DISSERTATIONSSCHRIFT

# Erfassung epistemologischer Überzeugungen von Lehramtsstudierenden

## Entwicklung und Validierung des StEB Inventars

zur
Erlangung des akademischen Doktorgrades
im Fach Erziehungswissenschaft
der Philosophischen Fakultät
an der Universität Passau

INKA S. HÄHNLEIN

2018

Erstgutachterin
Prof. Dr. Jutta Mägdefrau

Zweitgutachter
Prof. Dr. Johannes Hartig

*„Intellectual honesty consists in specifying precisely the conditions under which one is willing to give up one's position."*

Imre Lakatos

*Für meine Mutter*
*Kirstin Hähnlein*

# *Zusammenfassung*

**Erfassung epistemologischer Überzeugungen von Lehramtsstudierenden: Entwicklung und Validierung des StEB Inventars**

von Inka Sara HÄHNLEIN

Diese Arbeit befasst sich mit epistemologischen Überzeugungen von Lehramtsstudierenden in der Disziplin Erziehungswissenschaft. Das Konstrukt der epistemologischen Überzeugungen wird zunächst in einer historischen Synopse in seiner Entwicklung beleuchtet, wobei drei unterschiedliche Phasen der Konstruktgenese unterschieden werden können. Die zentralen Konzeptualisierungen epistemologischer Überzeugungen werden dargestellt. Aufbauend auf den jüngsten Erkenntnissen im Feld wird eine aktuelle Synthese vorgenommen, aus der sowohl wissenschaftlicher Anlass als auch praktischer Bedarf zu einer auf die Disziplin fokussierten Instrumententwicklung abgeleitet werden. Vor empirisch-methodologischem Hintergrund wird die klassische Testtheorie als geeignetes Methodenspektrum für die Instrumententwicklung herangezogen. Im Hinblick auf die inzwischen vorherrschende Forschungsauffassung, dass es sich bei den epistemologischen Überzeugungen um ein mehrdimensionales und disziplinspezifisches Konstrukt handelt, werden die aus dem Theoriegefüge gewonnenen Dimensionen zu einem Instrument, dem StEB (students' epistemological beliefs) Inventar, welches in drei aufeinanderfolgenden Studien entwickelt, geprüft, modifiziert und dessen Testwertinterpretationen validiert werden. Auf eine überregionale, formalisierte Expertenbefragung folgt eine zur EFA geeignete erste Untersuchung. Die darauf aufbauende Instrumentmodifikation wird in einer zweiten Stichprobe geprüft, die zugleich auch Validierungsstichprobe ist. Erneut werden die Modelle in dieser zweiten Stichprobe hinsichtlich Dimensionalität, Faktorladungen und Modellgüte mittels CFA und EFA getestet. Die Validierung folgt dem argumentationsbasierten Ansatz. Aus den Testwertinterpretationen leiten sich empirisch prüfbare Grundannahmen über den Testinhalt, die interne Teststruktur, die Beziehungen der StEB Testwerte mit Lernstrategien sowie angenommene Gruppenunterschiede ab. Die Prüfung der Grundannahmen stellt Evidenzen bereit, die die Gültigkeit der Testwertinterpetationen stützen. Das Resultat ist ein in Forschung und Praxis einsatzbereites, mehrdimensionales und disziplinspezifisches, quantitatives Erhebungsinstrument. Abschließend werden die forschungspragmatischen und praktischen Implikationen für das Feld der Erziehungswissenschaft und der Lehrerbildung diskutiert.

# *Danksagung*

Mein herzlicher Dank gilt meiner Erstbetreuerin Prof Dr. Jutta Mägdefrau für die Beratung, die Unterstützung und die Betreuung dieser Arbeit von der ersten Minute an. Ganz besonders möchte ich mich auch für die wertvollen inhaltlichen Rückmeldungen und Anregungen meines Zweitgutachters Prof. Dr. Johannes Hartig bedanken. Das Kolloquium „Methoden empirischer Bildungsforschung“ mit all seinen wunderbaren Teilnehmerinnen und Teilnehmern hat mich diesen Weg begleitet. Ich danke euch allen für eure stets konstruktive und wertschätzende Rückmeldung, eure Ratschläge, die Inspiration zum Wintersport und die rundum tolle Zeit.

Meinen lieben Passauer Kolleginnen Isolde Baumgartner, Martina Riel, Monika Wolff und meinem ehemaligen Kollegen Dr. Matthias Böhm gilt mein besonderer Dank. Euer aufbauender Zuspruch und der rege Austausch mit euch hat mir viel Energie für diese Arbeit gegeben. Vielen lieben Dank euch.

Ich möchte mich ausdrücklich bei den zehn Expertinnen und Experten für ihre Teilnahme an der Befragung und wertvolle Rückmeldung bedanken. Mein herzlicher Dank gilt außerdem den Hochschuldozierenden der Martin-Luther-Universität Halle-Wittenberg, der Pädagogischen Hochschule Freiburg sowie der Universitäten Passau, Bremen, Regensburg und Potsdam, die mich in der Datenerhebung unterstützt und diese in ihren Veranstaltungen ermöglicht haben. Insbesondere danke ich Dr. Thomas Lehmann und Philipp Streit für die Unterstützung.

Ohne den Rückhalt, den Mut und den Zuspruch aus meinem engsten Freundeskreis wäre diese Arbeit so nicht entstanden und mein Leben hätte mich auf einen ganz anderen Pfad geführt. Auch diesen Weg habt ihr mich mit Rat und Tat begleitet, wie schon so viele andere davor. Katha, Pablo, Dandan und Steffi, ich danke euch aus tiefstem Herzen dafür, dass ihr immer an mich glaubt und für mich da seid, euch mit mir freut, ärgert, traurig und ausgelassen seid.

# Inhaltsverzeichnis

**1 Einleitung** **1**
1.1 Ziel und Fragestellung . . . 6
1.2 Strukturierung . . . 8

**2 Epistemologische Überzeugungen** **11**
2.1 Forschungsentwicklung . . . 13
2.1.1 Erste Forschungsphase – Entwicklungsstufenmodelle . . . 13
2.1.2 Zweite Forschungsphase – Explizite Mehrdimensionalität . . . 17
2.1.3 Dritte Forschungsphase – Ausdifferenzierung . . . 19
2.2 Modelle epistemologischer Überzeugungen . . . 25
2.2.1 Scheme of Intellectual and Ethical Development von Perry . . . 25
2.2.2 Reflective Judgment Model von King und Kitchener . . . 31
2.2.3 Development of Epistemological Understanding von Kuhn . . . 35
2.2.4 Epistemological Belief System von Schommer . . . 37
2.2.5 Epistemological Theories von Hofer und Pintrich . . . 43
2.3 Entwicklung und Veränderung epistemologischer Überzeugungen . . . 51
2.3.1 Beschaffenheit der Entwicklung . . . 51
2.3.2 Auslöser von Entwicklung und Veränderung . . . 56
2.4 Explizite Mehrdimensionalität . . . 64
2.5 Domänenspezifität und Kontextabhängigkeit . . . 68
2.6 Theory of Integrated Domains in Epistemology (TIDE) Framework . . . 75
2.7 Verortung epistemologischer Überzeugungen in Theorien der Kognition und Metakognition . . . 80
2.8 Der Einfluss epistemologischer Überzeugungen auf den Lehr-Lernprozess . . . 85
2.8.1 Modelle zum Einfluss epistemologischer Überzeugungen auf den Lehr-Lernprozess . . . 87
2.8.2 Epistemologische Überzeugungen beim Lernen . . . 90
2.8.3 Epistemologische Überzeugungen beim Lehren . . . 110
2.9 Erhebung epistemologischer Überzeugungen . . . 113
2.9.1 Qualitative Erhebung . . . 113
2.9.2 Quantitative Erhebung . . . 114
2.9.3 Probleme der bestehenden quantitativen Erhebungsinstrumente . . . 122
2.10 Synthese . . . 126

**3 Zentrale Zielsetzung, Fragestellung und Grundannahmen** **135**

**4 Methodische Basis der Instrumententwicklung 143**
4.1 Testgütekriterien . . . 146
4.2 Argumentationsbasierter Validierungsansatz . . . 150
4.3 Prüfung des Testentwurfs und Validierung mittels EFA und CFA . . . 153
4.3.1 Exploratorische Faktorenanalyse . . . 153
4.3.2 Konfirmatorische Faktorenanalyse . . . 160

**5 Entwurf des StEB Inventars 169**
5.1 Dimensionen des StEB . . . 170
5.2 Itemformulierung . . . 174
5.3 Expertenbefragung . . . 179

**6 Exploratorische Modellbildung (Studie 1) 183**
6.1 Durchführung . . . 183
6.2 Datenbereinigung und deskriptive Datenanalyse . . . 185
6.2.1 Deskriptive Analyse der StEB-Items . . . 188
6.2.2 Analyse fehlender Werte . . . 191
6.3 Faktorenanalytische Auswertung der ersten Stichprobe . . . 193
6.3.1 Prüfung der Voraussetzung für die EFA . . . 193
6.3.2 Vorbereitung der EFA . . . 194
6.3.3 Ergebnisse der EFA . . . 197
6.3.4 Itemreduktion und inhaltliche Interpretation der Faktoren . . . 200
6.3.5 Prüfung der Faktorstruktur nach Itemreduktion . . . 204
6.4 Psychometrische Eigenschaften der Subskalen . . . 206
6.4.1 Itemanalyse . . . 206
6.4.2 Skalenanalyse . . . 209
6.5 Diskussion und Rückbezug zum theoretischen Hintergrund . . . 211

**7 Konfirmatorische Modellbildung und Kreuzvalidierung (Studie 2) 217**
7.1 Durchführung . . . 217
7.2 Inventar zur Erfassung von Lernstrategien im Studium . . . 219
7.3 Datenbereinigung . . . 221
7.4 Deskriptive Datenanalyse . . . 221
7.4.1 Deskriptive Analyse der StEB-Items . . . 225
7.4.2 Analyse fehlender Werte . . . 227
7.5 Faktorenanalytische Auswertung der Validierungsstichprobe . . . 229
7.5.1 Prüfung der Voraussetzung für die CFA . . . 229
7.5.2 Ergebnisse der CFA . . . 235
7.5.3 Zusammenfassung und Diskussion der Ergebnisse . . . 258
7.6 Item- und Skalenanalyse . . . 263
7.6.1 Messmodellprüfung . . . 263
7.6.2 Skalenanalyse . . . 266
7.6.3 Bestimmung der Testwerte des StEB . . . 268

7.6.4 Itemanalyse . . . . . 269

**8 Validierung der Testwertinterpretationen des StEB 273**
8.1 Testinhalt und interne Struktur des StEB . . . . . 275
8.2 Zusammenhänge zwischen den StEB-Dimensionen . . . . . 278
8.3 Beziehungen zwischen den StEB- und LIST-R-Testwerten . . . . . 281
8.3.1 Bereinigung der Daten und deskriptive Analyse des LIST-R . . . . . 282
8.3.2 Reliabilitätsanalyse der LIST-R-Skalen . . . . . 285
8.3.3 CFA des LIST-R . . . . . 287
8.3.4 Analyse der Zusammenhänge zwischen StEB und LIST-R . . . . . 289
8.4 Unterschiedsprüfungen mittels MANOVA . . . . . 294
8.5 Zusammenfassende Beurteilung der Evidenzen und Testwertinterpretationen 314

**9 Diskussion 317**
9.1 Zusammenfassung . . . . . 317
9.2 Bezug des StEB zu den Entwicklungsstufen epistemologischer Überzeugungen 328
9.3 Normative Sichtweise auf epistemologische Überzeugungen . . . . . 330
9.4 Einschränkungen . . . . . 333
9.5 Forschungspragmatische und praktische Implikationen des StEB Inventars . 338

**A Eingesetzte StEB Testentwürfe 341**

**B Itemwortlaut der StEB-Versionen 369**

**C Lernstrategien im Studium (LIST-R) 389**

**D CFA-Modelle zur Studie 2 393**

**E Parameterschätzungen zur Studie 2 401**

**F Itemkorrelations- und Residualkovarianzmatrizen Studie 2 419**

**G Deskriptive Daten zur Kreuzvalidierung 435**

**H Modell-Fit der Messmodellprüfungen 437**

**I Testwertvalidierung 441**

**Literatur 452**

# Abbildungsverzeichnis

2.1 Epistemologische Entwicklung nach dem *Perry Scheme* (siehe Perry, 1981, S. 80), in eigener Darstellung . . . . . 29
2.2 Gemeinsamkeiten in den Dimensionen epistemologischer Überzeugungen zwischen verschiedenen Modellen nach Hofer und Pintrich (1997, S. 113 ff.), in eigener Darstellung . . . . . 44
2.3 Zeitliche Entwicklung der epistemologischen Überzeugungen in verschiedenen Modellen von Wildenger, Hofer und Burr (2010, S. 227), in eigener Darstellung . . . . . 52
2.4 Mechanismen der Veränderung im *Integrative Model for Personal Epistemology* (IM) von Bendixen und Rule (2004) mit *personal epistemology multiplier* (Rule, 2003) nach Rule und Bendixen (2010, S. 98), in eigener Darstellung . . . . . 58
2.5 Vergleich der angenommenen Dimensionen epistemologischer Überzeugungen zwischen verschiedenen Modellen . . . . . 64
2.6 Überblick über die *Theory of Integrated Domains in Epistemology* (TIDE Framework) nach Muis, Bendixen und Haerle (2006, S. 30), in eigener Darstellung . 76
2.7 Epistemologische Prozesse (*epistemic metacognition*) integriert als Erweiterung der Metakognition nach Hofer (2004a, S. 49), in eigener Darstellung . . . . . 83
2.8 Einfluss epistemologischer Überzeugungen auf den Lernprozess nach Hofer (2001, S. 372), in eigener Darstellung . . . . . 87
2.9 *Embedded Systemic Model* nach Schommer-Aikins (2004, S. 25), in eigener Darstellung . . . . . 88
2.10 Darstellung der Lernstrategieklassifikation des MSLQ von Pintrich, Smith, Garcia und McKeachie (1991, S. 194), in eigener Darstellung . . . . . 98
2.11 Pfadmodell zum Einfluss epistemologischer Überzeugungen auf *approaches to learning* und die akademische Leistung (Cano, 2005, S. 214) in eigener Darstellung . . . . . 101
2.12 Pfadmodell zum Einfluss epistemologischer Überzeugungen auf Lernstrategien, Verstehen, kritisches Denken, Reflexion, Habitual Action und Lernleistung von Phan (2006) . . . . . 104
2.13 Bedeutende englischsprachige Fragebögen zur mehrdimensionalen, disziplinübergreifenden Erfassung epistemologischer Überzeugungen (überarbeitet und in eigener Darstellung nach Buehl, 2008, S. 70 ff.) . . . . . 116
2.14 Bedeutende englischsprachige Fragebögen zur mehrdimensionalen, disziplinspezifischen Erfassung epistemologischer Überzeugungen (überarbeitet und in eigener Darstellung nach Buehl, 2008, S. 70 ff.) . . . . . 118

2.15 Bedeutende deutschsprachige Fragebögen zur mehrdimensionalen, disziplinübergreifenden Erfassung epistemologischer Überzeugungen . . . . . 120
2.16 Bedeutende deutschsprachige Fragebögen zur mehrdimensionalen, disziplinspezifischen Erfassung epistemologischer Überzeugungen . . . . . 121
2.17 Zusammenhangsstruktur zwischen sophistizierten und wenig sophistizierten epistemologischen Überzeugungen und Lernstrategien . . . . . 129
2.18 Zusammenhangsstruktur zwischen den sophistizierten Ausprägungen epistemologischer Überzeugungen und Lernstrategien . . . . . 130

4.1 Darstellung eines CFA-Modells mit zwei latenten Variablen ($\xi$ und $\eta$) und sechs Items ($X_1$ bis $X_6$) . . . . . 161
4.2 Darstellung verschiedener Stufen der Äquivalenz von Messungen nach Bühner (2011, S. 400), in eigener Darstellung . . . . . 166

6.1 Kombinierte Darstellung verschiedener Tests zur Bestimmung der geeigneten Faktorenanzahl . . . . . 195
6.2 Histogramm der Residuen der 5-Faktorlösung . . . . . 197
6.3 Kombinierte Darstellung verschiedener Tests zur Bestimmung der geeigneten Faktorenanzahl . . . . . 204
6.4 Verortung der vier StEB-Dimensionen in den übergeordneten Bereichen epistemologischer Überzeugungen . . . . . 213

7.1 Streudiagramm für die Überzeugung zur Absolutheit des Wissens und Alter 230
7.2 Streudiagramm für die Überzeugung zur Absolutheit des Wissens und Fachsemester . . . . . 231
7.3 Streudiagramm für die Überzeugung zur Absolutheit des Wissens und Hochschulsemester . . . . . 231
7.4 Rekursives vierfaktorielles Gesamtmodell . . . . . 235
7.5 Standardisiertes vierfaktorielles Gesamtmodell ($N$ = 915) . . . . . 238
7.6 Geschätztes vierfaktorielles Gesamtmodell ($N$ = 915) . . . . . 240
7.7 Standardisiertes vierfaktorielles Modell zu Teilstichprobe S1 ($N$ = 457) . . . . 247
7.8 Standardisiertes vierfaktorielles Modell zu Teilstichprobe S2 ($N$ = 458) . . . . 247
7.9 Finales unstandardisiertes 26-Item Modell zu Teilstichprobe S1 ($N$ = 457) . . . 250
7.10 Finales unstandardisiertes 26-Item Modell zu Teilstichprobe S2 ($N$ = 458) . . . 251
7.11 Finales standardisiertes 26-Item Modell der Gesamtstichprobe ($N$ = 915) . . . 252
7.12 Finales unstandardisiertes 26-Item Modell der Gesamtstichprobe ($N$ = 915) . . 253

8.1 Empirische Zusammenhangsstruktur zwischen den StEB-Subskalen und Lernstrategien . . . . . 292
8.2 Durchschnittliche Ausprägung in der Überzeugung zur Absolutheit des Wissens nach Hochschule . . . . . 297
8.3 Boxplot der individuellen Skalenmittelwerte zur Absolutheit des Wissens nach Hochschule . . . . . 298

8.4 Durchschnittliche Ausprägung in der Überzeugung zur Einfachheit des Wissens nach Hochschule . . . 299
8.5 Boxplot der individuellen Skalenmittelwerte zur Einfachheit des Wissens nach Hochschule . . . 299
8.6 Durchschnittliche Ausprägung in der Überzeugung zur Multimodalität des Wissenserwerbs nach Hochschule . . . 300
8.7 Boxplot der individuellen Skalenmittelwerte zur Multimodalität des Wissenserwerbs nach Hochschule . . . 301
8.8 Durchschnittliche Ausprägung in der Überzeugung zur Absolutheit des Wissens nach Schulart . . . 304
8.9 Boxplot der individuellen Skalenmittelwerte zur Absolutheit des Wissens nach Schulart . . . 305
8.10 Durchschnittliche Ausprägung in der Überzeugung zur Einfachheit des Wissens nach Schulart . . . 306
8.11 Boxplot der individuellen Skalenmittelwerte zur Einfachheit des Wissens nach Schulart . . . 306
8.12 Durchschnittliche Ausprägung in der Überzeugung zur Einfachheit des Wissens nach fachlicher Ausrichtung . . . 312
8.13 Boxplot der individuellen Skalenmittelwerte zur Einfachheit des Wissens nach fachlicher Ausrichtung . . . 313

D.1 Initiales, unstandardisiertes, einfaktorielles Modell des Faktors Absolutheit ($N$ = 915) . . . 393
D.2 Initiales, standardisiertes, einfaktorielles Modell des Faktors Absolutheit ($N$ = 915) . . . 393
D.3 Initiales, unstandardisiertes, einfaktorielles Modell des Faktors Einfachheit ($N$ = 915) . . . 394
D.4 Initiales, standardisiertes, einfaktorielles Modell des Faktors Einfachheit ($N$ = 915) . . . 394
D.5 Initiales, unstandardisiertes, einfaktorielles Modell des Faktors Multimodalität ($N$ = 915) . . . 394
D.6 Initiales, standardisiertes, einfaktorielles Modell des Faktors Multimodalität ($N$ = 915) . . . 395
D.7 Initiales, unstandardisiertes, einfaktorielles Modell des Faktors Entwicklung ($N$ = 915) . . . 395
D.8 Initiales, standardisiertes, einfaktorielles Modell des Faktors Entwicklung ($N$ = 915) . . . 395
D.9 Standardisiertes vierfaktorielles Modell ($N$ = 915) . . . 396
D.10 Initiales, unstandardisiertes vierfaktorielles Modell zu Teilstichprobe 1 ($N$ = 457) . . . 397
D.11 Initiales, unstandardisiertes vierfaktorielles Modell zu Teilstichprobe 2 ($N$ = 458) . . . 397

D.12 Finales, standardisiertes einfaktorielles Modell des Faktors Entwicklung ($N$ = 915) . . . . . 398
D.13 Finales, standardisiertes einfaktorielles Modell des Faktors Entwicklung für Teilstichprobe 1 ($N$ = 457) . . . . . 398
D.14 Finales, standardisiertes einfaktorielles Modell des Faktors Entwicklung für Teilstichprobe 2 ($N$ = 458) . . . . . 398
D.15 Finales, standardisiertes 26-Item Modell Teilstichprobe 1 ($N$ = 457) . . . . . 399
D.16 Finales, standardisiertes 26-Item Modell Teilstichprobe 2 ($N$ = 458) . . . . . 399

I.1 Histogramm der individuellen Skalenmittelwerte zur Absolutheit des Wissens 442
I.2 Histogramm der individuellen Skalenmittelwerte zur Entwicklung des Wissens 442
I.3 Histogramm der individuellen Skalenmittelwerte zur Multimodalität des Wissenserwerbs . . . . . 443
I.4 Histogramm der individuellen Skalenmittelwerte zur Einfachheit des Wissens 443
I.5 Histogramm der individuellen Faktorwerte zur Absolutheit des Wissens . . . 444
I.6 Histogramm der individuellen Faktorwerte zur Entwicklung des Wissens . . 444
I.7 Histogramm der individuellen Faktorwerte zur Multimodalität des Wissenserwerbs . . . . . 445
I.8 Histogramm der individuellen Faktorwerte zur Einfachheit des Wissens . . . 445
I.9 Durchschnittliche Ausprägung in der Überzeugung zur Entwicklung des Wissens nach Hochschule . . . . . 446
I.10 Boxplot der individuellen Skalenmittelwerte zur Entwicklung des Wissens nach Hochschule . . . . . 446
I.11 Durchschnittliche Ausprägung in der Überzeugung zur Multimodalität des Wissenserwerbs nach Schulart . . . . . 447
I.12 Boxplot der individuellen Skalenmittelwerte zur Multimodalität des Wissenserwerbs nach Schulart . . . . . 447
I.13 Durchschnittliche Ausprägung in der Überzeugung zur Entwicklung des Wissens nach Schulart . . . . . 448
I.14 Boxplot der individuellen Skalenmittelwerte zur Entwicklung des Wissens nach Schulart . . . . . 448
I.15 Durchschnittliche Ausprägung in der Überzeugung zur Entwicklung des Wissens nach fachlicher Ausrichtung . . . . . 449
I.16 Boxplot der individuellen Skalenmittelwerte zur Entwicklung des Wissens nach fachlicher Ausrichtung . . . . . 449
I.17 Durchschnittliche Ausprägung in der Überzeugung zur Absolutheit des Wissens nach fachlicher Ausrichtung . . . . . 450
I.18 Boxplot der individuellen Skalenmittelwerte zur Absolutheit des Wissens nach fachlicher Ausrichtung . . . . . 450
I.19 Durchschnittliche Ausprägung in der Überzeugung zur Multimodalität des Wissenserwerbs nach fachlicher Ausrichtung . . . . . 451

I.20 Boxplot der individuellen Skalenmittelwerte zur Multimodalität des Wissenserwerbs nach fachlicher Ausrichtung . . . . . 451

# Tabellenverzeichnis

2.1 Stufen epistemologischen Verstehens nach Kuhn und Pearsall (2000, S. 311) . 36
2.2 Zuordnung der 12 Subsets zu den 5 Subskalen des SEQ . . . . . 40
2.3 Verortung des epistemologischen Verstehens in der kognitiven Verarbeitung bei Kitchener (1983) und Kuhn und Pearsall (2000) nach Hofer (2001, S. 364), in eigener Darstellung . . . . . 82
2.4 Bivariate Korrelationen zwischen epistemologischen Überzeugungen, Motivation, Lernstrategien und Selbstkonzept nach Urhahne und Hopf (2004, S. 79) . . . . . 100
2.5 Bivariate Korrelationen zwischen epistemologischen Überzeugungen, Motivation, Lernstrategien und Selbstkonzept nach Urhahne und Harms (2006, S. 193) . . . . . 102
2.6 Bivariate Korrelationen zwischen epistemologischen Überzeugungen und Lernstrategien (Dahl, Bals & Turi, 2005, S. 265) . . . . . 105
2.7 Multiple Regressionen epistemologischer Überzeugungen auf Lernstrategien (Dahl et al., 2005, S. 267 ff.) – Teil 1 . . . . . 106
2.8 Multiple Regressionen epistemologischer Überzeugungen auf Lernstrategien (Dahl et al., 2005, S. 267 ff.) – Teil 2 . . . . . 106
2.9 Bivariate Korrelationen zwischen epistemologischen Überzeugungen und Lernstrategien (vgl. Schworm & Gruber, 2017, S. 153) . . . . . 108

4.1 Stichprobengröße für EFA in Abhängigkeit von Kommunalität für ML-Methode nach Bühner (2011, S. 345), in eigener Darstellung . . . . . 156
4.2 Eignung der Daten nach dem KMO- und MSA-Maß nach Kaiser und Rice (1974) 157
4.3 Richtlinien für die Bewertung der Modellpassung nach den Fit-Indizes . . . . 165

5.1 Item-Skalen-Zuordnung des 126-Item StEB Inventar Testentwurfs . . . . . 176
5.2 Ausprägungen epistemologischer Überzeugungen je Skala für die Pole des 126-Item StEB Inventar Testentwurfs . . . . . 177
5.3 Zusammengefasste Expertenurteile über die Subskalen des 126-Item StEB Testentwurfs . . . . . 181
5.4 Item-Skalen-Zuordnung des 98-Item StEB Inventar Testentwurfs . . . . . 182

6.1 Häufigkeiten der studierten Schularten ($N$ = 217) . . . . . 185
6.2 Häufigkeiten und prozentuale Anteile nach Studienbeginn und Fachsemester ($N$ = 217) . . . . . 186

6.3 Häufigkeiten und prozentuale Anteile nach Studienbeginn und Hochschulsemester ($N$ = 217) . . . 186
6.4 Häufigkeiten und prozentuale Anteile der studierten Erst- und Zweitfächer ($N$ = 217) . . . 187
6.5 Häufigkeiten und prozentuale Anteile der studierten Dritt- und Viertfächer ($N$ = 217) . . . 187
6.6 Deskriptive Statistik der StEB-Items (Studie 1) – Teil 1 . . . 189
6.7 Deskriptive Statistik der StEB-Items (Studie 1) – Teil 2 . . . 190
6.8 Mustermatrix mit Faktorladungen $\lambda$ und zugehörigen Kommunalitäten der 5-Faktorlösung (89-Item StEB Inventar Version) – Teil 1 . . . 198
6.9 Mustermatrix mit Faktorladungen $\lambda$ und zugehörigen Kommunalitäten der 5-Faktorlösung (89-Item StEB Inventar Version) – Teil 2 . . . 199
6.10 Mustermatrix mit Faktorladungen $\lambda$ und zugehörigen Kommunalitäten und Uniqueness der 34-Item 4-Faktorlösung ($N$ = 217) . . . 205
6.11 Itemschwierigkeiten $P_i$ und korrigierte Itemtrennschärfen $r_{ix}$ der 34-Item 4-Faktorlösung . . . 208
6.12 Skalenhomogenität und interne Konsistenz der Skalen . . . 209
6.13 Korrelationen zwischen den Faktoren nach dem Ergebnis der EFA der 34-Item StEB-Version . . . 210
6.14 Item-Skalen-Zuordnung des 34-Item StEB Inventars . . . 212

7.1 Aufbau des LIST-R mit Item-Skalen-Zuordnung, Itemanzahl je Skala und Reliabilität $\alpha$ . . . 219
7.2 Häufigkeiten und prozentuale Anteile nach Hochschule (nach Ausschluss der Ausreißer) . . . 222
7.3 Häufigkeiten und prozentuale Anteile nach studierter Schulart (nach Ausschluss der Ausreißer) . . . 222
7.4 Häufigkeiten und prozentuale Anteile nach Studienbeginn und Fachsemester (nach Ausschluss der Ausreißer) . . . 223
7.5 Häufigkeiten und prozentuale Anteile der studierten Erst- und Zweitfächer (nach Ausschluss der Ausreißer) . . . 224
7.6 Häufigkeiten und prozentuale Anteile der studierten Dritt- und Viertfächer (nach Ausschluss der Ausreißer) . . . 224
7.7 Häufigkeiten der Fächer und Zuordnung der Fächer zu Bereichen ($N$ = 915) . . . 225
7.8 Deskriptive Statistik der StEB-Items (nach Ausschluss der Ausreißer) . . . 226
7.9 Schiefe und Exzess der StEB-Items . . . 233
7.10 Geschätzte semipartielle Regressionsgewichte (Ladungen $\lambda_{i1}$) der ersten Modellspezifikation ($N$ = 915) . . . 237
7.11 Geschätzte quadrierte Korrelationen der Faktoren mit Items der ersten Modellspezifikation ($N$ = 915) . . . 239
7.12 Geschätzte, unstandardisierte Itemfehlervarianzen ($N$ = 915) . . . 241
7.13 Modell-Fit (robust) der vier initialen einfaktoriellen Modelle ($N$ = 915) . . . 243

7.14 Mustermatrix der vierfaktoriellen Hauptachsenanalyse mit oblimin-Rotation, Itemkommunalitäten $h^2$ und Itemtrennschärfen $r_{ix}$ ($N$ = 915) . . . . . 245
7.15 Initialer Modell-Fit (robust) der vierfaktoriellen Modelle für Teilstichprobe S1 und S2 . . . . . 248
7.16 Modell-Fit (robust) für die schrittweise modifizierten Modelle der Teilstichprobe S1 und S2 . . . . . 249
7.17 Modell-Fit (robust) der finalen vierfaktoriellen Modelle der Gesamtstichprobe (GS) und Teilstichprobe S1 und S2 . . . . . 253
7.18 Modell-Fit (robust) für die finalen einfaktoriellen Modelle der Gesamtstichprobe und Teilstichprobe S1 und S2 . . . . . 256
7.19 Modell-Fit (robust) für die finalen einfaktoriellen Modelle mit korrelierten Fehlervarianzen der Gesamtstichprobe und Teilstichprobe 1 und 2 . . . . . 257
7.20 Item-Skalen-Zuordnung des finalen 26-Item StEB Inventars . . . . . 261
7.21 Messmodellprüfung des finalen vierfaktoriellen 26-Item Modells ($N$ = 915) . 265
7.22 Skalenreliabilität der finalen Subskalen des StEB . . . . . 267
7.23 Skalenhomogenität der finalen StEB-Subskalen getrennt nach Stichproben . . 267
7.24 Itemschwierigkeit getrennt nach Stichproben . . . . . 270
7.25 Korrigierte Itemtrennschärfe $r_{ix}$ der 26 verbleibenden StEB-Items je Stichprobe 271

8.1 Kovarianzen zwischen den Faktoren nach dem Ergebnis der CFA der 26-Item StEB-Version . . . . . 278
8.2 Deskriptive Analyse der LIST-R-Items für die Konstruktvalidierung (bereinigte Stichprobe, $N$ = 891) – Teil 1 . . . . . 283
8.3 Deskriptive Analyse der LIST-R-Items für die Konstruktvalidierung (bereinigte Stichprobe, $N$ = 891) – Teil 2 . . . . . 284
8.4 Skalenhomogenität und interne Konsistenz der LIST-R-Skalen im Vergleich Wild (2011) . . . . . 285
8.5 Part-whole korrigierte Itemtrennschärfe der LIST-R-Items ($N = 891$) . . . . . 286
8.6 Modell-Fit der einfaktoriellen LIST-R-Modelle und des neunfaktoriellen Gesamtmodells . . . . . 288
8.7 Korrelationskoeffizienten (Pearson) zwischen den Skalenmittelwerten des StEB und den Skalensummenwerten des LIST-R . . . . . 290
8.8 Mittelwerte und Standardabweichungen der StEB-Überzeugungsausprägungen getrennt nach Hochschule . . . . . 295
8.9 Mittelwerte und Standardabweichungen der StEB Überzeugungsausprägungen getrennt nach Schulart . . . . . 302
8.10 Mittelwerte und Standardabweichungen der StEB Überzeugungsausprägungen getrennt nach Studienfortschritt . . . . . 307
8.11 Zuordnung der Fächer zu Bereichen und Häufigkeiten der Fächer ($N$ = 849) . 309
8.12 Mittelwerte und Standardabweichungen der StEB-Überzeugungsausprägungen getrennt nach fachlicher Ausrichtung . . . . . 311

B.1 Itemwortlaut, -polung und -ursprung für den Itempool – Teil 1 . . . . . 369

B.2 Itemwortlaut, -polung und -ursprung für den Itempool – Teil 2 . . . . . 370
B.3 Itemwortlaut, -polung und -ursprung für den Itempool – Teil 3 . . . . . 371
B.4 Itemwortlaut, -polung und -ursprung für den Itempool – Teil 4 . . . . . 372
B.5 Itemwortlaut, -polung und -ursprung für den Itempool – Teil 5 . . . . . 373
B.6 Itemwortlaut, -polung und -ursprung für den Itempool – Teil 6 . . . . . 374
B.7 Itemwortlaut, -polung und -ursprung für den Itempool – Teil 7 . . . . . 375
B.8 Itemwortlaut, -polung und -ursprung der 98-Item StEB-Version – Teil 1 . . . . 376
B.9 Itemwortlaut, -polung und -ursprung der 98-Item StEB-Version – Teil 2 . . . . 377
B.10 Itemwortlaut, -polung und -ursprung der 98-Item StEB-Version – Teil 3 . . . . 378
B.11 Itemwortlaut, -polung und -ursprung der 98-Item StEB-Version – Teil 4 . . . . 379
B.12 Itemwortlaut, -polung und -ursprung der 98-Item StEB-Version – Teil 5 . . . . 380
B.13 Itemwortlaut, -polung und -ursprung der 98-Item StEB-Version – Teil 6 . . . . 381
B.14 Code, Wortlaut und Polung der Items der 34-Item StEB-Version – Teil 1 . . . . 382
B.15 Code, Wortlaut und Polung der Items der 34-Item StEB-Version – Teil 2 . . . . 383
B.16 Code, Wortlaut und Skalenzugehörigkeit der Items des finalen StEB Inventars – Teil 1 . . . . . 384
B.17 Code, Wortlaut und Skalenzugehörigkeit der Items des finalen StEB Inventars – Teil 2 . . . . . 385
B.18 Auf bestehenden Fragebögen basierende Items der 34-Item StEB-Version . . . 386
B.19 Auf bestehenden Fragebögen basierende Items der finalen 26-Item StEB-Version 387

C.1 Itemwortlaut Lernstrategien im Studium (LIST-R, Wild (2011)) - Teil 1 . . . . 390
C.2 Itemwortlaut Lernstrategien im Studium (LIST-R, Wild (2011)) - Teil 2 . . . . 391

E.1 Geschätzte semipartielle Regressionsgewichte (Ladungen $\lambda_{i1}$) der Items des finalen 26-Item vierfaktoriellen Modells mit korrelierten Itemfehlervarianzen ($N$ = 915) . . . . . 401
E.2 Geschätzte Kovarianzen zwischen den Items und Faktoren des finalen 26-Item vierfaktoriellen Modells mit korrelierten Itemfehlervarianzen ($N$ = 915) . 402
E.3 Geschätzte $R^2$ der Items des finalen 26-Item vierfaktoriellen Modells mit korrelierten Itemfehlervarianzen ($N$ = 915) . . . . . 402
E.4 Geschätzte Intercepts (Itemmittelwerte) der Items des finalen 26-Item vierfaktoriellen Modells mit korrelierten Itemfehlervarianzen ($N$ = 915) . . . . . 403
E.5 Geschätzte Itemfehlervarianzen und Faktorvarianzen des finalen 26-Item vierfaktoriellen Modells mit korrelierten Itemfehlervarianzen ($N$ = 915) . . . . . 404
E.6 Korrelierte Fehlervarianzen des vierfaktoriellen 26-Item Gesamtmodells sowie der vier einfaktoriellen Modelle der Gesamtstichprobe ($N$ = 915) . . . . . 405
E.7 Geschätzte semipartielle Regressionsgewichte (Ladungen $\lambda_{i1}$) der Items des finalen 26-Item vierfaktoriellen Modells ($N$ = 915) . . . . . 406
E.8 Geschätzte Kovarianzen zwischen den Faktoren des finalen 26-Item vierfaktoriellen Modells ($N$ = 915) . . . . . 407
E.9 Geschätzte $R^2$ der Items des finalen 26-Item vierfaktoriellen Modells ($N$ = 915) 407

E.10 Geschätzte Intercepts (Itemmittelwerte) der Items des finalen 26-Item vierfaktoriellen Modells ($N$ = 915) . . . . . 408
E.11 Geschätzte Itemfehlervarianzen und Faktorvarianzen des finalen 26-Item vierfaktoriellen Modells ($N$ = 915) . . . . . 409
E.12 Geschätzte semipartielle Regressionsgewichte (Ladungen $\lambda_{i1}$) der Items des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe S1 ($N$ = 457) . . . . 410
E.13 Geschätzte Kovarianzen zwischen den Faktoren des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe S1 ($N$ = 457) . . . . . 411
E.14 Geschätzte $R^2$ der Items des finalen 26-Item vierfaktoriellen Modells von Teilstichprobe S1 ($N$ = 457) . . . . . 411
E.15 Geschätzte Intercepts (Itemmittelwerte) der Items des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe S1 ($N$ = 457) . . . . . 412
E.16 Geschätzte Itemfehlervarianzen und Faktorvarianzen des finalen 26-Item vierfaktoriellen Modells von Teilstichprobe S1 ($N$ = 457) . . . . . 413
E.17 Geschätzte semipartielle Regressionsgewichte (Ladungen $\lambda_{i1}$) der Items des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe S2 ($N$ = 458) . . . . 414
E.18 Geschätzte Kovarianzen zwischen den Faktoren des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe S2 ($N$ = 458) . . . . . 415
E.19 Geschätzte $R^2$ der Items des finalen 26-Item vierfaktoriellen Modells von Teilstichprobe S2 ($N$ = 458) . . . . . 415
E.20 Geschätzte Intercepts (Itemmittelwerte) der Items des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe S2 ($N$ = 458) . . . . . 416
E.21 Geschätzte Itemfehlervarianzen und Faktorvarianzen des finalen 26-Item vierfaktoriellen Modells von Teilstichprobe S2 ($N$ = 458) . . . . . 417

F.1 Itemkorrelationsmatrix der Gesamtstichprobe ($N$ = 915) – Teil 1 . . . . . 420
F.2 Itemkorrelationsmatrix der Gesamtstichprobe ($N$ = 915) – Teil 2 . . . . . 421
F.3 Itemkorrelationsmatrix der Teilstichprobe S1 ($N$ = 457) – Teil 1 . . . . . 422
F.4 Itemkorrelationsmatrix der Teilstichprobe S1 ($N$ = 457) – Teil 2 . . . . . 423
F.5 Itemkorrelationsmatrix der Teilstichprobe S2 ($N$ = 458) – Teil 1 . . . . . 424
F.6 Itemkorrelationsmatrix der Teilstichprobe S2 ($N$ = 458) – Teil 2 . . . . . 425
F.7 Geschätzte Residualkovarianzmatrix des initialen 34-Item vierfaktoriellen Modells der Gesamtstichprobe ($N$ = 915) – Teil 1 . . . . . 427
F.8 Geschätzte Residualkovarianzmatrix des initialen 34-Item vierfaktoriellen Modells der Gesamtstichprobe ($N$ = 915) – Teil 2 . . . . . 428
F.9 Geschätzte Residualkovarianzmatrix des finalen vierfaktoriellen Modells der Teilstichprobe 1 ($N$ = 457) – Teil 1 . . . . . 429
F.10 Geschätzte Residualkovarianzmatrix des finalen vierfaktoriellen Modells der Teilstichprobe 1 ($N$ = 457) – Teil 2 . . . . . 430
F.11 Geschätzte Residualkovarianzmatrix des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe 2 ($N$ = 458) – Teil 1 . . . . . 431

F.12 Geschätzte Residualkovarianzmatrix des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe 2 ($N$ = 458) – Teil 2 . . . . . 432
F.13 Geschätzte Residualkovarianzmatrix des finalen 26-Item vierfaktoriellen Modells der Gesamtstichprobe ($N$ = 915) – Teil 1 . . . . . 433
F.14 Geschätzte Residualkovarianzmatrix des finalen 26-Item vierfaktoriellen Modells der Gesamtstichprobe ($N$ = 915) – Teil 2 . . . . . 434

G.1 Deskriptive Statistik der StEB-Items von Teilstichprobe 1 ($N$ = 457) . . . . . 435
G.2 Deskriptive Statistik der StEB-Items von Teilstichprobe 2 ($N$ = 458) . . . . . 436

H.1 Modell-Fit der Messmodellprüfung für die einfaktoriellen Modelle der Gesamtstichprobe ($N$ = 915) . . . . . 438
H.2 Modell-Fit der Messmodellprüfung der einfaktoriellen Modelle von Teilstichprobe 1 ($N$ = 457) . . . . . 439
H.3 Modell-Fit der Messmodellprüfung der einfaktoriellen Modelle von Teilstichprobe 2 ($N$ = 458) . . . . . 440

I.1 Deskriptive Analyse der StEB-Items für Prüfung der Grundannahme 4 (bereinigte Stichprobe, $N$ = 891) . . . . . 441

# 1 Einleitung

Während die Epistemologie[1] als Bereich der Philosophie klassischerweise die Frage nach der Natur des menschlichen Wissens, seinen Ursprüngen, Grenzen, Methoden und Begründungen betrifft (Audi, 1999, S. 273), befasst sich die Psychologie und pädagogische Psychologie seit über fünf Jahrzehnten mit dem Konstrukt der individuellen, subjektiven Überzeugungen zu Wissen und zum Wissensprozess. Diese Forschung betrifft Fragen danach, wie Individuen Konzeptionen zu Wissen und dem Wissensprozess entwickeln, wie diese Konzeptionen beschaffen sind und wie sie genutzt werden, um aus Erfahrungen Bedeutung zu konstruieren (Hofer, 2002, S. 4). Die individuellen Überzeugungen über die Beschaffenheit des Wissens und den Prozess des Wissens einer Person werden als *epistemologische Überzeugungen*, *epistemische Überzeugungen*, *epistemische Kognition* oder auch als *persönliche Epistemologie* bezeichnet (Hofer, 2016). Die Überzeugungen zur *Natur des Wissens* (*nature of knowledge*) betreffen individuelle Vorstellungen darüber, wie Wissen selbst beschaffen und organisiert ist. Die Überzeugungen zur *Natur des Wissensprozesses*[2] (*nature of knowing*) betreffen die individuellen Vorstellungen darüber, wie der Prozess des Wissens abläuft sowie Vorstellungen über die Herkunft von Wissen und Begründbarkeit von Wissensurteilen. Im Zentrum der Beschaffenheit des Wissensprozesses steht vor allem die Frage: Wann *weiß* man etwas? (Hofer, 2004a).

Die Bezeichnung *epistemologische Überzeugungen* – neben oben genannten anderen Bezeichnungen – wird verwendet, um „a set of mental processes that involve the development and employment of one's conceptions of knowledge and knowing" (Hofer, 2016, S. 20) zu beschreiben. Es handelt sich um diejenigen mentalen Prozesse einer Person, die die Entwicklung sowie den kognitiven und metakognitiven Einsatz der eigenen Konzeptionen zu Wissen und zum Wissensprozess umfassen (vgl. Bendixen & Rule, 2004; Hofer, 2004a, 2005, 2016; Muis et al., 2006). Die wissensbezogenen Überzeugungen entstehen in dynamischer Interaktion des Individuums mit dem Umfeld (Baxter Magolda, 2004; Belenky, Clinchy, Goldberger & Tarule, 1986; Bendixen & Rule, 2004; Hofer & Pintrich, 1997) durch Erziehung, Bildung, persönliche Erfahrung und kulturelle Einflüsse und entwickeln sich über die Lebensspanne weiter (z.B. Bendixen & Rule, 2004; Blignaut, Mokwinski, Paechter, Rebmann & Sulimma, 2010; Dann, 1994; Hofer, 2004a; Mokwinski, 2010; Muis et al., 2006; Müller, 2009). Sie werden im instruktionalen, akademischen und sozio-kulturellen Kontext aktiviert

[1] Der Begriff *Episteme* stammt aus dem Griechischen und bezeichnet nach Aristoteles *Wissen* und zwar Wissen, das sich auf Begründungen bezieht (Wissen warum, nicht Wissen was) (Audi, 1999, S. 46), während *logos Erklärung* bezeichnet (Audi, 1999, S. 273).

[2] Wissensprozess stellt hier eine Übersetzung des englischen Begriffs *knowing* dar. Im Englischen wird zwischen *beliefs about knowing und knowledge* unterschieden (Hofer & Pintrich, 1997). Beide Begriffe würden im Deutschen mit *Wissen* übersetzt werden. Da diese Übersetzung missverständlich wäre, wird *knowing* in dieser Arbeit als Wissensprozess bezeichnet, während *knowledge* mit Wissen übersetzt wird.

und von diesem beeinflusst (Hofer, 2005, 2006b; Muis et al., 2006). Bildung beeinflusst die epistemologische Entwicklung, indem es die Fertigkeiten von Individuen stärkt, Informationen kritisch zu beurteilen sowie miteinander in Konflikt stehende Annahmen, Theorien und Evidenzen zu koordinieren (King & Kitchener, 1994; Kuhn, 1991).

Das Konstrukt epistemologischer Überzeugungen besteht, zurückgehend auf Schommer (1990), aus mehreren Dimensionen. Die angenommene mehrdimensionale Struktur konnte mittlerweile in zahlreichen Studien gefunden werden (z.B. Bråten & Olaussen, 2005; Conley, Pintrich, Vekiri & Harrison, 2004; Jehng, Johnson & Anderson, 1993; Muis et al., 2006; Schraw, Bendixen & Dunkle, 2002). Da es viele unterschiedliche Konzeptionen gibt, hängt es vom zugrundegelegten Modell ab, welche Überzeugungsdimensionen angenommen werden. Die folgenden vier Dimensionen konnten jedoch als modellübergreifende Kerndimensionen des Konstrukts identifiziert werden (Hofer & Pintrich, 1997): Überzeugungen zur *Quelle des Wissens*, *Rechtfertigung des Wissens*, *Sicherheit des Wissens* und *Einfachheit des Wissens*. Dieses System von Überzeugungen ist nach (Hofer & Pintrich, 1997) theorieähnlich organisiert (vgl. Hofer, 2004a, 2005, 2016). Über diese vier hinaus konnte eine Dimension zu den Überzeugungen zur *Entwicklung des Wissens* identifiziert werden (Conley et al., 2004). Die Überzeugungen zur Stabilität, Struktur und Entwicklung des Wissens können dem übergeordneten Bereich *Beschaffenheit des Wissens* zugeordnet werden, während die Überzeugungen zur Quelle und Rechtfertigung des Wissens zum Bereich *Beschaffenheit des Wissensprozesses* gehören.

Mittlerweile hat sich im Forschungsfeld der epistemologischen Überzeugungen zudem die Annahme durchgesetzt, dass es nicht nur verschiedene Dimensionen epistemologischer Überzeugungen gibt, die bei Personen unterschiedlich ausgeprägt sein können, sondern dass diese auch in Bezug auf unterschiedliche Fach- und Wissensbereiche (Domänen) verschieden ausgeprägt sein können. Es wird angenommen, dass Individuen sowohl über 1. allgemeine, übergreifende nicht-akademische Überzeugungen (Alltag, Arbeitsumfeld etc.), 2. allgemeine akademische Überzeugungen als auch über 3. disziplinspezifische akademische Überzeugungen (Hofer, 2005; Muis et al., 2006) verfügen. Es handelt sich also um drei wissensbezogene Überzeugungssysteme, die jedoch untereinander verbunden sind. Mit Eintritt in das Bildungssystem und mit fortschreitenden Bildungserfahrungen entwickelt ein Individuum in Auseinandersetzung mit verschiedenen Unterrichtsfächern (in der Schule) und akademischen Disziplinen (in der Universität) parallel jeweils verschiedene disziplinspezifische epistemologische Überzeugungen (*Theory of Integrated Domains in Epistemology Framework*, TIDE, Muis et al., 2006). Mit zunehmender akademischer Erfahrung und Spezialisierung in einer bestimmten Disziplin nimmt der Einfluss dieser disziplinspezifischen wissensbezogenen Überzeugungen auf die allgemeinen akademischen Überzeugungen zu (ebd.). Jedes der drei Überzeugungssysteme umfasst jeweils alle Überzeugungsdimensionen (ebd.). In vielen Studien konnten sowohl intraindividuelle als auch interindividuelle fachdisziplinspezifische epistemologische Überzeugungen bezüglich verschiedener Überzeugungsdimensionen gefunden werden (z.B. Buehl, Alexander & Murphy, 2002; Estes, Chandler, Horvath & Backus, 2003; Hofer, 2000; Jehng et al., 1993; Schommer & Walker,

1995; Youn, 2000). So existieren beispielsweise interindividuelle Unterschiede zwischen Studierenden unterschiedlicher Fachrichtungen (Studienfächer) in Bezug auf Geistes- , Sozialwissenschaft und Naturwissenschaft (z.B. Jehng et al., 1993; Lampert, 1990; Schoenfeld, 1992; Youn, 2000). Intraindividuelle Unterschiede zeigten sich in Abhängigkeit vom jeweiligen Wissensbereich (z.B. Buehl et al., 2002; Estes et al., 2003; Hofer, 2000; Schommer & Walker, 1995).

Jede dieser Überzeugungsdimensionen kann als Kontinuum mit zwei Polen beschrieben werden (Hofer & Pintrich, 1997). Das Kontinuum reicht von weniger entwickelten, weniger sophistizierten wissensbezogenen Überzeugungen bis hin zu sophistizierten, höher entwickelten wissensbezogenen Überzeugungen. Jede der Überzeugungsdimensionen entwickelt sich demnach in Richtung einer sophistizierteren Perspektive auf Wissen und des Prozesses des Wissens weiter (Bendixen & Rule, 2004; Muis et al., 2006; Rule & Bendixen, 2010). In der weniger sophistizierten Perspektive wird Wissen als feststehend, unveränderlich, aus einfachen, unverbundenen Fakten bestehend, von Autoritäten vermittelt und als zweifelsfrei wahr oder falsch klassifizierbar betrachtet, ohne einer kritischen Auseinandersetzung mit den Wissensquellen zu bedürfen (Bendixen & Rule, 2004; Hofer & Pintrich, 1997; Muis et al., 2006). In der sophistizierteren Ausprägung wird Wissen als vorläufig, veränderlich, komplex miteinander verbunden, persönlich in Interaktion mit dem sozialen Umfeld konstruiert und als kritisch zu hinterfragen und zu rechtfertigen betrachtet (Bendixen & Rule, 2004; Hofer & Pintrich, 1997; Muis et al., 2006).

Die Überzeugungen einer Person darüber, was Wissen ist und wie man etwas weiß (Hofer & Bendixen, 2012, S. 227), beeinflussen, wie diese Person neuen Informationen begegnet und diese verarbeitet sowie ihre Verstehens- Schlussfolgerungs-, Problemlöse- und ganz generell Lernprozesse (Hofer & Bendixen, 2012). Sie haben eine handlungsleitende und -steuernde Funktion (Dann, 1994; Helmke, 2005; Mokwinski, 2010 und beeinflussen beispielsweise das Ausmaß, zu dem Individuen die Bedeutungskonstruktion beim komplexen Problemlösen gelingt sowie die Intensität der Auseinandersetzung (Hofer & Pintrich, 2002) und demnach die Herangehensweise an Lernaufgaben (Hofer & Sinatra, 2010, S. 116). Ein sophistizierteres epistemologisches Überzeugungssystem wirkt sich förderlich auf multiple lernrelevante Faktoren und demnach förderlich auf den Lernprozess eines Individuums aus, während weniger fortgeschrittene Überzeugungen eher lernhemmend sind (Hofer, 2001; Schommer-Aikins, 2002). Sophistizierte wissensbezogene Überzeugungen wirken sich positiv auf alle Phasen des selbstgesteuerten Lernens aus (Bromme, 2010). Sie wirken als Linse, durch die Lernende die Lernaufgabe wahrnehmen, Lerninhalte erfassen und dienen der erfolgreichen Anpassung des Lernverhaltens (z.B. bezüglich Lernstrategienauswahl) an die wahrgenommenen Anforderungen und Rahmenbedigungen der Aufgabe (Bromme, 2010; Pieschl, Stahl & Bromme, 2013). Sie sind lernförderlich im Hinblick auf die akademische Leistung (z.B. Barnard, Lan, Crooks & Paton, 2008; Hofer, 2000; Köller, Baumert & Neubrand, 2000; Schommer-Aikins & Hutter, 2002; Trautwein, Lüdke & Beyer, 2004), Lernemotionen (z.B. Muis, Psaradellis, Lajoie, Di Leo & Chevrier, 2015), Motivation (z.B. Buehl & Alexander, 2005; Hofer, 1999), Selbstkonzept (z.B. Urhahne & Harms, 2006), Fachinteresse (z.B. Köller et al., 2000), Selbststeuerungsfähigkeit (z.B. Bromme, 2010) sowie Kontroll- und

Selbstwirksamkeitsüberzeugungen (z.B. Muis & Foy, 2007; Paulsen & Feldman, 1999). Sie führen zu förderlichem Lernverhalten, indem zur Lernaufgabe passendere Lernstrategien gewählt und eingesetzt werden (z.B. Bromme, 2010; Hofer, 2001; Schommer-Aikins, 2002) und unterstützen den Lerntransfer (Jacobson & Spiro, 1995), den *conceptual change* (z.B. Franco et al., 2012; Mason & Boscolo, 2004; Qian & Alvermann, 1995), die Informationssuche und -bewertung (z.B. Mason, Boldrin & Ariasi, 2010b; Pieschl et al., 2013; Whitmire, 2004) und das Testverarbeiten und -verstehen (z.B.Bendixen, 2002; Bråten & Strømsø, 2004; Kardash & Noel, 2000; Schommer-Aikins, 2004).

Auch auf den Lehrprozess wirken sich die wissensbezogenen Überzeugungen entscheidend aus (Hofer, 2001). Sie haben einen Einfluss auf das unterrichtliche Verhalten von Lehrenden gegenüber Lernenden. Sie beeinflussen, „wie der Lehrende unterrichtliche Situationen auffasst, welche Handlungsmöglichkeiten er in Betracht zieht und welche er letztlich auswählt, wie er Effekte seines Eingreifens bewertet und wie er nachträglich sein gesamtes Handeln begründet oder rechtfertigt“ (Wahl, 1979, S. 209). Für praktizierende Lehrkräfte haben epistemologische Überzeugungen deshalb eine besondere Relevanz, da das komplexe Problemlösen, im Sinne des Managements von komplexen Lernumgebungen, die multiple Stakeholder umfassen, im Kern des Lehrberufs liegt (Brownlee, Ferguson & Ryan, 2017) und das Gelingen des Problemlösens von den eigenen wissensbezogenen Überzeugungen beeinflusst wird (Hofer & Pintrich, 2002). Es konnte bereits empirisch gezeigt werden, dass die epistemologischen Überzeugungen von praktizierenden Lehrkräften ihre Lehrtätigkeit, ihre Unterrichtsstrategien und ihre Erwartungen an die Schülerinnen und Schüler beeinflussen (Brownlee, Schraw & Berthelsen, 2011; Buehl & Fives, 2009).

Die vielfältigen Einflüsse der epistemologischen Überzeugungen auf das Lernen betreffen natürlich auch Lehramtsstudierende, von denen, wie von allen Studierenden, erwartet wird selbstgesteuert für ihr Studium zu lernen, d.h. Verantwortung für ihren eigenen Lernprozess zu übernehmen (Schworm & Gruber, 2017, S. 143). Für Lehramtsstudierende sind die eigenen epistemologischen Überzeugungen insofern von besonderer Bedeutung, als dass sie sowohl das eigene Lernen im Studium als auch deren Konzeption von Lernkontexten für das Lernen von Schülerinnen und Schülern beeinflussen (Buehl & Fives, 2016, S. 248). Die Lehrerbildung erstreckt sich über einen langen Zeitraum und umfasst in Deutschland das Studium, Praktika und das Referendariat. Angehende Lehrkräfte müssen also kontinuierlich selbstgesteuert lernen – und zwar sowohl in den verschiedenen Wissensdomänen ihrer Unterrichtsfächer als auch im erziehungswissenschaftlichen Wissensbereich – um ihr professionelles Handeln weiterzuentwickeln und das Lernen von Schülerinnen und Schülern erfolgreich fördern zu können (Buehl & Fives, 2016). Es konnte bereits empirisch gezeigt werden, dass die wissensbezogenen Überzeugungen von Lehramtsstudierenden beeinflussen, wie tief ihr erworbenes Verständnis der Inhalte ihrer Lehrveranstaltungen (*teacher education courses*) geht (Buehl & Fives, 2016; Yadav & Koehler, 2007). Sie filtern, wie Lehramtsstudierende das Lernen in ihren universitären Veranstaltungen – die ja ihres Studiums zu Lehrkräften dienen – erleben und inwiefern sie sinnvolle, tiefenorientierte Lernansätze betreiben (Many, Howard & Hoge, 2002; Muis, 2004; Peng & Fitzgerald, 2006; Yadav & Koehler, 2007). Auch die selbstständige Recherche und das selbstständige Lesen und

Erarbeiten von Texten, die wichtige Aspekte eines Studiums darstellen, sind vom epistemologischen Überzeugungssystem abhängig. Sophistizierte epistemologische Überzeugungen beeinflussen das Verhalten von Lernenden bei der selbstregulierten Informationssuche und -bewertung positiv (Hofer, 2004b; Mason et al., 2010b; Pieschl et al., 2013; Whitmire, 2004). Je fortgeschrittener das epistemologische Überzeugungssystem ist, desto eher wenden Studierende viel Zeit und Sorgfalt bei der Planung der Informationssuche auf, nutzen multiple Quellen, differenzieren zwischen der Herkunft von Inhalten und vergleichen kritisch zwischen verschiedenen Inhalten. Sophistizierte epistemologische Überzeugungen wirken sich förderlich auf die Textverarbeitung und das Textverstehen aus (z.B. Bråten & Strømsø, 2004; Kardash & Noel, 2000; Schommer, Crouse & Rhodes, 1992; Schommer & Walker, 1995; Schommer-Aikins, 2004; Schraw et al., 2002), weshalb Lehramtsstudierende mit fortgeschrittenem wissensbezogenen Überzeugungssystem besser und erfolgreicher mit Studientexten lernen können. Schon bei angehenden Lehrerinnen und Lehrern stehen die individuellen wissensbezogenen Überzeugungen in starker Verbindung mit der unterrichtlichen Zielsetzung und den im Unterricht umgesetzten Handlungen (z.B. im Rahmen von Praktika oder des Referendariats) (z.B. Kang, 2008; Yadav & Koehler, 2007). Sie wirken sich auf ihre Entscheidungsfindung, Planung, Durchführung und Bewertung der praktischen Anwendung des Gelernten aus (Buehl & Fives, 2016; Yadav & Koehler, 2007). Angehende Lehrkräfte mit sophistizierteren epistemologischen Überzeugungen nutzen eher pädagogische Praktiken, die die Kinder ermutigen unabhängig zu denken und Wissen zu konstruieren, als solche mit weniger sophistizierten epistemologischen Überzeugungen (Brownlee, Schraw, Walker & Ryan, 2016).

Für Lehramtsstudierende ist es folglich von großer Bedeutung, über sophistizierte epistemologische Überzeugungen zu verfügen, „that can influence the depth of understanding they take from teacher education courses, the development of deep and varied knowledge systems for teaching, and the use of effective, informed decision-making across a range of contexts“ (Brownlee et al., 2011, S. 303). Da epistemologische Überzeugungen aber disziplinspezifisch sind (Muis et al., 2006), wäre es erstrebenswert, dass Lehramtsstudierende sowohl bezüglich der Wissensdomänen ihrer jeweiligen Unterrichtsfächer als auch bezüglich des erziehungswissenschaftlichen Wissens über sophistizierte Überzeugungsysteme verfügen.

Aufgrund der besonderen Bedeutung der epistemologischen Überzeugungen von Lehramtsstudierenden für ihr eigenes Lernen und die spätere Handlungspraxis, sollte eine Förderung sophistizierter epistemologischer Überzeugungen – und zwar sowohl für die Fachdisziplin(en) als auch die Erziehungswissenschaft – in der Lehrerbildung ein zentrales Ziel darstellen (vgl. Brownlee et al., 2011, 2016). Um fortgeschrittene epistemologische Überzeugungen bei Lehramtsstudierenden zielgerichtet fördern zu können, ist es unumgänglich, das vorherrschende epistemologische Überzeugungssystem zunächst zu verstehen (Chai, Khine & Teo, 2006) und dazu muss es erfasst werden können. Nur wenn die aktuell vorherrschenden wissensbezogenen Überzeugungen ermittelt werden können, lässt sich feststellen, bei wem und in Bezug auf welche Überzeugungsdimensionen Förderungsbedarf vorliegt. Und nur wenn die epistemologischen Überzeugungen erhoben werden können, kann der

Erfolg einer durchgeführten Förderungsmaßnahme ermittelt werden. Unter solchen Fragestellungen, welche die Gesamtheit der angehenden Lehrkräfte betreffen, eignen sich quantitative Erhebungssinstrumente besser als qualitative, da sie ökonomisch einsetzbar und auszuwerten sind, sich gut für die Erfassung von Veränderungen in Merkmalsausprägungen eignen sowie die einfache Untersuchung von Verbindungen zu anderen lernrelevanten Konstrukten ermöglichen (Mason, 2016).

Für die verschiedenen Fachdisziplinen wie beispielsweise Mathematik, Physik und Geschichte ist die quantitative Erhebung mittels Fragebogen bereits möglich, da es passende Instrumente gibt. Es gibt jedoch bisher kein quantitatives Erhebungsinstrument, das sich explizit für die Erfassung der Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen und Wissensprozess eignet. Und das, obwohl die Erfassung dieser epistemologischen Überzeugungen von großer theoretischer und praktischer Relevanz ist. Dieses erziehungswissenschaftliche Wissen umfasst schließlich pädagogische und didaktische Inhalte, die für den späteren Lehrberuf zentral sind, wie z.B. Unterrichtsplanung, Klassenführung, Umgang mit Unterrichtsstörungen und Störungsprävention, Umgang mit Heterogenität, Variablen der Unterrichtsqualität, Unterrichtsmethoden, -prinzipien und Sozialformen. Und wie oben gezeigt wurde, beeinflusst das epistemologische Überzeugungssystem eines Lehramtsstudierenden zu diesem erziehungswissenschaftlichen Wissen, wie tief diese Inhalte im Studium verstanden werden, was davon gelernt wird und was davon wie in die Unterrichtspraxis überführt wird (Buehl & Fives, 2016; Yadav & Koehler, 2007). Demnach ist es sowohl aus Forschungsperspektive als auch für die Anwendung notwendig, ein quantitatives Erhebungsinstrument zu entwickeln, das es ermöglicht, die epistemologischen Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen empirisch zugänglich zu machen.

## 1.1 Ziel und Fragestellung

Das übergeordnete Ziel dieser Arbeit ist demnach die Entwicklung eines Instruments, genannt StEB (*students' epistemological beliefs*) Inventar, das die disziplinspezifischen mehrdimensionalen epistemologischen Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen empirisch erfassbar macht.

Bei der Erhebung ist jedoch zu berücksichtigen, dass die epistemologischen Überzeugungen im Kontext aktiviert werden (Hofer, 2005, 2006b; Muis et al., 2006). Da sich die Studierenden im akademischen Kontext befinden und das Lernen im Studium immer in einem speziellen instruktionalen Kontext stattfindet, beeinflussen diese sehr stark, welche epistemologischen Überzeugungen in den Lehramtsstudierenden aktiviert werden und demnach auch erhoben werden können. Um sicherzustellen, dass die befragten Lehramtsstudierenden die „richtigen" epistemologischen Überzeugungen, d.h. diejenigen, die hier von Interesse sind – die Überzeugungen zum erziehungswissenschaftliche Wissen und Wissensprozess – aktivieren, wird der instruktionalen Kontext in dieser Instrumentenentwicklung eingrenzt. Ganz konkret wird sich die Erhebung auf die *epistemologischen Überzeugungen von*

*Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen und Wissensprozess aus Studientexten* beziehen. Dieser Teilbereich wurde gewählt, da das selbstständige Lesen und Lernen aus Texten im Sinne des selbstgesteuerten Lernens ein wichtiger Aspekt des erziehungswissenschaftlichen Lehramtsstudiums ist und, wie zuvor dargelegt, die disziplinspezifischen epistemologischen Überzeugungen die Textverarbeitung und das Textverstehen stark beeinflussen (z.B. Bråten & Strømsø, 2004; Kardash & Noel, 2000; Schommer et al., 1992; Schommer & Walker, 1995; Schommer-Aikins, 2004; Schraw et al., 2002).

Aus der zentralen Zielsetzung ergibt sich die zentrale Fragestellung, der im Rahmen dieser Arbeit nachgegangen wird. Es soll geprüft werden, ob sich in der Faktorstruktur des StEB die epistemologischen Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen aus Studientexten den theoretischen Annahmen ensprechend abbilden lassen. In diesem Zusammenhang soll überprüft werden, ob sich die aus der Theorie abgeleiteten Dimensionen epistemologischer Überzeugungen im StEB Inventar abbilden lassen, welche inhaltlichen Faktoren sich ergeben, ob die resultierende Struktur mit dem Theoriegefüge vereinbar ist und über welche psychometrischen Eigenschaften die finale Version des StEB Inventars verfügt. Im Rahmen der Validierung des StEB Inventars werden außerdem acht Grundannahmen untersucht, mittels derer sich die Gültigkeit der Interpretationen der Testergebnisse und Angemessenheit der intendierten Testnutzung prüfen lassen. In diesem Zusammenhang werden zum Beispiel auch die Beziehungen zwischen den StEB-Überzeugungsskalen und den Lernstrategieskalen des LIST-R (Lernstrategien im Studium revidiert, Wild, 2011) untersucht. Es wird außerdem geprüft, ob sich Unterschiede in den Ausprägungen epistemologischer Überzeugungen zwischen Lehramtsstudierenden verschiedener Hochschulstandorte und Fachsemester finden lassen.

In dieser Arbeit wird ein pädagogisch-psychologischer Ansatz epistemologischer Überzeugungen vertreten (Hofer & Pintrich, 1997). Durch die Verwendung von Items und Ratingskalen bietet diese Perspektive die Möglichkeit der ökonomischen Diagnostik epistemologischer Überzeugungen, sodass interpersonelle Vergleiche in der Ausprägung epistemologischer Überzeugungen möglich werden und Zusammenhänge mit anderen Variablen untersucht werden können (ebd.). Der methodologische Hintergrund dieser Instrumententwicklung ist die klassische Testtheorie. Sie gibt die Kriterien vor, auf denen diese Instrumententwicklung basiert und nach denen die Skalen- und Itementwicklung durchgeführt, validiert und analysiert wird. Folglich gliedert sich der Entwicklungsprozess übergeordnet in 1. die Entwicklung des Testentwurfs, 2. die empirische Prüfung des Testentwurfs und die 3. Validierung der Testwertinterpretationen (Bühner, 2011; Hartig, Frey & Jude, im Druck).

Diese drei Schritte stellen jedoch nur den groben Rahmen der Instrumententwicklung dar und beinhalten jeweils viele Teilschritte. Im Folgenden wird die Strukturierung dieser Arbeit detaillierter vorgestellt.

## 1.2 Strukturierung

Der *theoretische Hintergrund* bildet den Anfang einer Instrumententwicklung. Dieser wird in **Kapitel 2** ausführlich dargestellt. Zunächst wird ein Überblick über die Forschungsentwicklung in drei übergeordneten Phasen gegeben, um die großen, den Forschungsbereich bestimmenden Veränderungen und Forschungstrends vorzustellen. Im Anschluss werden fünf bedeutende Konzeptionen epistemologischer Überzeugungen vorgestellt, die einerseits für den gesamten Forschungsbereich, aber im speziellen auch für die Fokussierung dieser Arbeit von zentraler Bedeutung sind. Dem folgt eine Darstellung des Entwicklungsverlaufs epistemologischer Überzeugungen, die die Natur der Entwicklung sowie die Auslöser von Entwicklung und Veränderung beleuchtet. Danach werden mit der Mehrdimensionalität, Domänenspezifität, Kontextabhängigkeit und Verortung epistemologischer Überzeugungen in den Theorien der Kognition und Metakognition zentrale Forschungsdiskurse, bedeutende Modelle sowie wichtige Implikationen für die Definition, Betrachtung und Abgrenzung des Konstrukts dargestellt. In Abschnitt 2.8 schließt sich eine umfassende Darstellung der Beziehung epistemologischer Überzeugungen zum Lehr- und Lernprozess an. Hier wird anhand zweier Modelle und zahlreicher Forschungsbefunde aufgezeigt, welche Verbindungen die epistemologischen Überzeugungen zu verschiedenen lernrelevanten Faktoren aufweisen und wie sie den Lernprozess und das Lehren beeinflussen. Einen Schwerpunkt bildet dabei die Verbindung der wissensbezogenen Überzeugungen zu den Lernstrategien, weshalb hier auch eine bedeutende Lernstratgiekonzeption skizziert wird. Es folgt eine überblicksartige Darstellung der zentralen Erhebungsinstrumente epistemologischer Überzeugungen, wobei der Schwerpunkt dabei auf der quantitativen Erhebung liegt und Limitationen der bestehenden Skalen diskutiert werden. Den Abschluss des Theorieteils bildet die Synthese, die die Positionierung des Konstrukts epistemologischer Überzeugungen in dieser Arbeit darstellt, überblicksartig zusammenfasst, auf welche theoretischen Modelle und Konzeptionen welche Aspekte der Definition des Konstrukts in dieser Arbeit zurückgehen, die besondere Relevanz epistemologischer Überzeugungen für den Lehr-Lernprozess von angehenden Lehrkräften herausstellt sowie die Relevanz der Instrumententwicklung dieser Arbeit argumentativ herleitet.

Aus dieser Synthese des theoretischen Hintergrundes leitet sich in **Kapitel 3** die übergeordnete Zielsetzung dieser Arbeit – die Testentwicklung – ab, aus der sich zwei zentrale Fragestellungen ergeben. Die Zielsetzung lässt sich in spezifische Interpetationen der Testergebnisse, sogenannte Testwertinterpretationen, überführen. Daraus ergibt sich wiederum der konkrete Nutzen, der angestrebte Einsatz des zu entwickelnden Tests. Aus den spezifizierten Testwertinterpretationen werden sodann prüfbare Grundannahmen abgeleitet, die der späteren Validierung des Tests dienen.

**Kapitel 4** betrifft den *methodischen Hintergrund* der Instrumententwicklung nach der klassischen Testtheorie. Über die Erfassung von Konstrukten im Rahmen von psychologischen Tests wird zunächst zur klassischen Testtheorie (KTT) hingeleitet. Deren Grundanahmen sowie die verschiedenen Messmodelle werden im Anschluss knapp aufgeführt, bevor die Testgütekriterien Objektivität, Reliabilität und Validität beschrieben werden. Hier wird

bezüglich der Validierung einerseits auf das alte Validitätsverständnis, das noch verschiedene Validitätsarten unterscheidet, und danach ganz zentral auf den aktuellen argumentationsbasierten Ansatz eingegangen. Im Folgenden wird der idealtypische Ablauf einer Instrumententwicklung nach der klassischen Testtheorie unter Einsatz der faktorenanalytischen Verfahren skizziert, indem zunächst die Teilschritte der exploratorischen Faktorenanalyse und anschließend die der konfirmatorischen Faktorenanalyse aufgezeigt werden.

**Kapitel 5** befasst sich mit der *Entwicklung des Testentwurfes*. Im Rahmen dieses Kapitels wird zunächst knapp auf die Eingrenzung des Zielkonstrukts, die Definition des Messgegenstandes und die Konstruktionsstrategie des StEB eingegangen, bevor die eigentlichen Definitionen der theoretisch angenommenen fünf Dimensionen des StEB Inventars dargestellt werden. Dem schließt sich in einem eigenen Teilkapitel die Itemsformulierung an, die das Festlegen der Indikatoren und Zielgruppe, die Bestimmung des Testziels, das Bestimmen des Itemformats sowie die eigentliche Itemformulierung nach speziellen Formulierungsrichtlinien umfasst, mit der die konkrete Skalenkonstruktion umgesetzt wird. Dieses Kapitel schließt mit der Beschreibung der Durchführung und Auswertung der Expertenbefragung, mittels derer der Itempool geprüft und optimiert wird.

**Kapitel 6** widmet sich im Rahmen einer ersten Studie der *exploratorischen Modellbildung* des StEB Inventars. Hier wird aufgezeigt, wie der durch Expertenurteile optimierte StEB Testentwurf bei einer ersten Stichprobe eingesetzt und unter Einsatz der exploratorischen Faktorenanalyse (EFA) empirisch geprüft wird. Den Anfang des Kapitels bilden die Beschreibung der Durchführung, Datenbereinigung, deskriptiven Analyse und Analyse der fehlenden Werte. Dieser folgt eine ausführliche Darstellung der eigentlichen faktorenanalytischen Auswertung der Daten dieser ersten Stichprobe, die in einer Item- und Faktorreduktion, einer neuen Faktorstruktur und demnach einer Modifikation des StEB Inventars mündet. Die Faktorstruktur wird über eine erneute exploratorische Faktorenanalyse auf Stabilität geprüft, bevor mittels einer Item- und Skalenanalyse die psychometrischen Eigenschaften der Skalen ermittelt werden. Dem schließt sich die Diskussion der ersten Studie an, die einen Rückbezug zu den betreffenden Fragestellungen und der Theorie umfasst. Zum Abschluss dieses Kapitels wird das auf Basis der ersten Studie rekonstruierte StEB vorgestellt.

**Kapitel 7** umfasst im Rahmen einer zweiten Studie die *konfirmatorische Modellbildung* des StEB Inventars. Hier wird das auf Basis der ersten Stichprobe modifizierte StEB Inventar im Sinne des idealtypischen Ablaufs einer Instrumententwicklung nach der KTT im Rahmen einer zweiten Studie mit neuen, größeren Stichprobe geprüft und validiert (vgl. Brown, 2015; Bühner, 2011; Fabrigar, Wegener, MacCallum & Strahan, 1999; R. B. Kline, 2016, S. 198). Zum Einsatz kommt dazu der modellgenerierende Ansatz (vgl. Jöreskog, 1993) der konfirmatorischen Faktorenanalyse (CFA), da die CFA in dieser Art als ein Instrument zur Konstruktion und Evaluation von Skalen eingesetzt werden kann. Den Anfang bildet hier die Beschreibung der Durchführung und Datenerhebung. An dieser Stelle wird auch das Inventar zur Erfassung von Lernstrategien im Studium kurz vorgestellt (LIST-R, revidierte Version, Wild, 2011), da es neben dem StEB Inventar im Rahmen der Erhebung eingesetzt wurde. Dem schließen sich die Datenbereinigung und die deskriptive Analyse sowie die Analyse der

fehlenden Werte an. Im Rahmen der folgenden faktorenanalytischen Auswertungen mittels CFAs wird geprüft, ob sich die Faktorstruktur replizieren lässt, die aus der ersten Studie resultierte. Im Anschluss wird eine Modelloptimierung vorgenommen, die über eine Teilung der zweiten Stichprobe in zwei Substichproben Schritt für Schritt kreuzvalidiert wird. In einem eigenen Teilkapitel werden im Folgenden die Item- und Skalenanalyse dargestellt. Das Kapitel schließt mit einer Diskussion der gesamten faktorenanalytischen Auswertung der zweiten Studie.

**Kapitel 8** umfasst die Validierung der Testwertinterpretationen des StEB Inventars nach dem argumentationsbasierten Ansatz. Dazu werden die aus den Testwertinterpretationen abgeleiteten Grundanannahmen geprüft. Die Prüfung diente dazu, Evidenzen zu sammeln, die die Testwertinterpretationen stärken oder aber gegen sie sprechen, um schlussendlich eine Aussage darüber treffen zu können, ob die intendierten Testwertinterpretationen und folglich der geplante Einsatz des StEB zulässig sind. Als Evidenzquellen werden die interne Teststruktur, der Testinhalt und die Beziehung der StEB-Testwerte zu anderen Variablen herangezogen. In diesem Rahmen wird auch die Zusammenhangsstruktur zwischen den resultierenden StEB-Skalen und den Lernstrategieskalen des LIST-R (Wild, 2011) untersucht. Postulierte Unterschiede in den Überzeugungsdimensionen des finalen StEB Inventars zwischen den Lehramtsstudierenden in Abhängigkeit der Hochschule, des Fachsemesters und der fachlichen Ausrichtung werden mittels MANOVAs geprüft.

In **Kapitel 9** wird der gesamte Entwicklungsverlauf des StEB Inventars zunächst zusammengefasst dargestellt. Im Anschluss werden verschiedene kritische Aspekte diskutiert. Die Arbeit schließt mit einer Diskussion der formalen, forschungspragmatischen und praktischen Implikationen des StEB Inventars für das Feld der Erziehungswissenschaft und der darin angesiedelten Lehrerbildung.

# 2 Epistemologische Überzeugungen

Die Forschung zu den epistemologischen Überzeugungen geht auf verschiedene Forschungstraditionen und Disziplinen zurück und wird von der Entwicklungspsychologie, Bildungswissenschaft und pädagogischen Psychologie vorangetrieben (Hofer, 2002, S. 4). Auch Wissenschaftlerinnen und Wissenschaftler speziell im Bereich der tertiären Bildung, Lehrerbildung, der Leseforschung, den Naturwissenschaften und der Mathematik forschen zu diesem Konstrukt (ebd.). Als Folge davon haben die Theoriebildung und empirische Forschung in diesem Bereich stark zugenommen. Da ihre Ursprünge allerdings in den unterschiedlichsten Forschungsrichtungen und -interessen liegen, erschienen die Forschungsergebnisse sehr getrennt voneinander und verschiedene Begrifflichkeiten und ein uneinheitliches Verständnis des Konstrukts hat sich entwickelt (Hofer, 2002, S. 4). Den subjektiven Überzeugungen einer Person zum Wissen und Wissensprozess werden schon in der Psychologie je nach Konzeption ganz unterschiedliche Fachtermini zugeordnet (Hofer, 2016). Unter anderem wird es bezeichnet als: *epistemological position* (Scheme of Intellectual and Ethical Development, Perry, 1970, 1981), *epistemological standard* (M. P. Ryan, 1984), *epistemological perspectives* (Women's Ways of Knowing, Belenky et al., 1986), *Ways of Knowing, epistemological reflection* (Development of Epistemological Reflection, Baxter Magolda, 1992), *reflective thinking, reflective judgment* (Development of Reflective Thinking, King & Kitchener, 1994, 2002, 2004), *epistemological views* (Argumentative Reasoning, Kuhn, 1991), *epistemological understanding* (Development of Epistemological Understanding, Kuhn & Pearsall, 2000; Kuhn & Weinstock, 2002), *epistemological beliefs* (Epistemological Belief System, Schommer, 1990; Schommer-Aikins & Hutter, 2002), *epistemic beliefs* (Schraw et al., 2002), *personal epistemology, epistemological theories* (Epistemological Theories, Hofer, 2000, 2001, 2002; Hofer & Pintrich, 1997), *epistemological resources* (Epistemological Resources, Hammer & Elby, 2002; Louca, Elby, Hammer & Kagey, 2004), *epistemic metacognition* (Hofer, 2004a), *epistemological beliefs* (Beliefs about Academic Knowledge, Buehl & Alexander, 2001, 2006; Buehl et al., 2002) und *epistemic cognition* (Brownlee, Ferguson & Ryan, 2017; Chinn, Bruckland & Samarapungavan, 2011; Greene, Muis & Pieschl, 2010; Kitchener, 1983).

Das deutet schon an, dass es sich auch innerhalb der Psychologie nicht um ein einheitlich definiertes Konstrukt handelt, sondern viele unterschiedliche theoretische Konzeptionen existieren (vgl. Hofer, 2016; Hofer & Pintrich, 2002). Auch die Methoden der Erhebung sind

sehr verschieden (Mason, 2016). Es wurde in den vergangen Jahren bereits zweimal der Versuch unternommen, sich auf einen gemeinsamen Überbegriff, der all diese Forschungsbeiträge umfasst, zu einigen und diesen einzuführen. Ein einheitlicher Begriff, mit dem auf diesen Forschungsinhalt referenziert werden kann, würde es Forschern aus den unterschiedlichen Teilbereichen erleichtern, ähnliche Forschungsstränge zu identifizieren und voneinander zu lernen (Hofer, 2016). 2002 wurde im Rahmen der ersten überarbeiteten Anthologie (Hofer, 2002) der Überbegriff persönliche Epistemologie (*personal epistemology*) gewählt. 2016 brachten Greene, Sandoval und Braten ein Handbuch heraus mit dem Titel *Handbook of Epistemic Cognition* (Greene, Sandoval & Bråten, 2016). Im Rahmen dieses Werkes wird der Überbegriff *epistemische Kognition* (*epistemic cognition*) vorgeschlagen sowie die Bedeutung und der Umfang dieses Konstrukts definiert (Hofer, 2016). Es wird jedoch explizit hervorgehoben, dass diese neue Konzeptspezifikation nicht impliziert, dass andere Konzeptionen damit ihre Gültigkeit verlieren oder gar ersetzt werden sollen (Hofer, 2016, S. 21). Es ist mit Blick auf die zahlreichen aktuellen Forschungsbeiträge festzustellen, dass sich weder der Begriff *personal epistemology* noch der Begriff *epistemic cognition* einheitlich durchsetzen konnte und neben diesen beiden nach wie vor sehr häufig die Bezeichnungen *epistemological beliefs* oder auch *epistemic belief* verwendet werden. Auch innerhalb von Sammelbänden ist die Begrifflichkeit nicht einheitlich (z.B. Bendixen & Feucht, 2010; Bernholt, Gruber & Moschner, 2017; Brownlee et al., 2011; Elen, Stahl, Bromme & Clarebout, 2011; Fives & Gill, 2014; Hofer & Pintrich, 2002; Khine, 2008; Schraw, Brownlee, Olafson & VanderVeldt, 2017) und das trifft sogar für die Sammelbände zu, die doch eigentlich postulieren, eine einheitliche Begrifflichkeit einführen zu wollen (Greene et al., 2016; Hofer & Pintrich, 2002). Parallel dazu sind auch in deutschen Forschungsbeiträgen, Sammelbänden und Dissertationen die Begriffe *epistemologische Überzeugungen* (z.B. Berding & Lamping, 2014; Klopp, 2014; Müller, 2009; Peter, 2015; Priemer, 2006; Trautwein et al., 2004; Urhahne & Harms, 2006; Zinn, 2013) sowie *epistemische Überzeugungen* (z.B. Bernholt et al., 2017; Hartinger, Kleickmann & Hawelka, 2006; Merk, 2016) sehr gebräuchlich.

Im Rahmen dieser Arbeit wird der im deutschen sehr gebräuchliche Begriff der *epistemologischen Überzeugungen* verwendet. Als Synonyme dazu werden außerdem auch die Bezeichnungen *Überzeugungen zum Wissen und Wissensprozess*, *epistemologisches Überzeugungssystem* oder aber *wissensbezogene Überzeugungen* eingesetzt.

Die Darstellung des theoretischen Hintergrundes zu den epistemologischen Überzeugungen beginnt im Folgenden mit Beschreibung der Forschungsentwicklung in drei übergeordneten Phasen, um einen Überblick über die großen, den Forschungsbereich bestimmenden Veränderungen und Forschungstrends zu geben.

## 2.1 Forschungsentwicklung

Die Forschungsentwicklung des psychologischen Konstrukts *epistemologische Überzeugung* – das aktuell auch häufig epistemische Kognition (*epistemic cognition*) oder persönliche Epistemologie (*personal epistemology*) genannt wird (Greene et al., 2016) – werden in den Werken von Hofer und Pintrich (1997), Hofer (2016) und Greene, Azevedo und Torney-Purta (2008) detailliert dargestellt. Die Darstellung hier dient dazu, einen Überblick über die Entwicklungen zu geben, die die Forschung zu epistemologischen Überzeugungen seit ihren Anfängen geprägt haben.

Die Forschung zu epistemologischen Überzeugungen erfuhr seit ihren Anfängen mit den Theorien kognitiver Entwicklung von Piaget (1950, 1954, 1970), den Studien[1] von Perry zur intellektuellen Entwicklung, sowie den psycho-epistemologischen Profilen von Royce (1978) und Royce und Mos (1980) einen starken Aufschwung (Greene et al., 2016). Die Forschungsentwicklung und mit ihr die zahlreichen unterschiedlichen Konzeptionen zu epistemologischen Überzeugungen lassen sich nach Hofer (2016) in drei Forschungswellen unterteilen. Jede dieser Forschungswellen wird gekennzeichnet durch zentrale Entwicklungen und besonders einflussreiche Forschungsbeiträge.

### 2.1.1 Erste Forschungsphase – Entwicklungsstufenmodelle

Die ersten bedeutenden Modelle epistemologischer Überzeugungen waren, angefangen mit dem *Perry Scheme* von Perry (1970), Entwicklungsstufenmodelle, die alle im angloamerikanischen Raum entstanden. Ziel von Perrys qualitativer phänomenologischer Längsschnittstudie war ursprünglich, die unterschiedlichen Arten und Weisen zu erklären, wie Studierende sich die Bedeutung von Lernerfahrungen erklären. Das *Perry Scheme* beschreibt neun Positionen intellektueller und ethischer Entwicklung, die den folgenden vier übergeordneten Positionen zugeordnet werden: Dualismus (*dualism*), Multiplizismus (*multiplizism*), kontextueller Relativismus (*contextual relativism*) und Selbstverpflichtung im Relativismus (*commitment within relativism*) (Perry, 1970). Auch wenn Perry selbst im Zusammenhang mit seinem Stufenmodell nicht von epistemologischer sondern von intellektueller und ethischer Entwicklung sprach, beschreibt sein Modell ein sich weiterentwickelndes Verständnis über das Wissen, den Wissensprozess und Lernen (Moore, 2002). Es wird deshalb im Forschungsbereich übereinstimmend als das erste epistemologische Entwicklungsstufenmodell bezeichnet und stellt die Grundlage für weitere Modelle dar. Aufgrund seiner Relevanz für alle weiteren Instrumententwicklungen im Forschungsfeld wird es im Kapitel 2.2.1 (ab S. 25) im Detail dargestellt.

Auch wenn sich diese Forschungsphase durch die Entwicklung der verschiedenen Entwicklungsstufenmodelle charakterisieren lässt, gab es doch auch zwei bedeutende andere Ansätze von Royce und Kollegen sowie M. P. Ryan (1984). Zeitgleich zu Perry entwickelten Royce und Kollegen (im US-amerikanischen Raum) im Rahmen ihrer empirischen Forschung zum individuellen Wissensprozess (*how people know*) die psycho-epistemologischen Profile (Royce, 1959, 1967, 1974, 1978, 1983; Royce & Mos, 1980; Royce & Smith, 1964), die

[1] Perry begann 1953 mit seinen Pilotstudien.

allerdings weit weniger Bekanntheit im Forschungsfeld der epistemologischen Überzeugungen erlangten. Beim Ansatz von Royce handelt es sich, anders als bei den meisten übrigen Modellen epistemologischer Überzeugungen dieser Forschungsphase, nicht um ein Entwicklungsstufenmodell. Royce setzt die epistemologischen Überzeugungen explizit in Verbindung mit den kognitiven Prozessen des Wissenserwerbs, indem er Wissen als diejenigen Kognitionen der kognitiven Struktur eines Individuums definiert, die epistemologisch gerechtfertigt werden können (Royce, 1978).

Ein weiterer wichtiger Forschungsbeitrag dieser Forschungsphase stammt von M. P. Ryan (1984). Er nutzte selbst eine quantitative Erhebungsmethode und empfahl diese als objektivere Erhebungsmethode. M. P. Ryan (1984) betrachtete im Rahmen seiner Forschung zur metakognitiven Überwachung des Textverstehens bei Studierenden einen einzelnen Aspekt von Perrys Forschung, die Überzeugungen zur Struktur des Wissens. Er erfasste, ob die Studierenden entweder über dualistische oder aber relativistische Überzeugungen zu Wissen verfügten. Das betrifft Überzeugungen darüber, dass Wissen entweder richtig oder falsch sein muss (dualistisch) oder aber kontextabhängig (relativistisch) ist. Zur Klassifikation der Studierenden in Dualisten bzw. Relativisten ließ M. P. Ryan (1984) sie Aussagen zu Überzeugungen von Perry (1968) auf einer Likertskala einschätzen. Der Forschungsbeitrag von M. P. Ryan (1984) liegt einerseits darin, dass er zeigte, dass sich quantitative Erhebungsmethoden für die Erfassung epistemologischer Überzeugungen eignen. Die quantitative Erfassung, die er als objektivere Methode empfahl, machte es auch möglich, die Verbindung zwischen epistemologischen Überzeugungen und Lernprozessen zu untersuchen (Schommer-Aikins, 2004, S. 20 f.).

Die Forschung zu den *Women's Ways of Knowing* von Belenky et al. (1986), die auf dem Interviewprotokoll von Gilligan (1982) für moralische Entwicklung basiert, stellt das zweite Entwicklungsstufenmodell dar. Nicht nur erweiterten sie den Forschungsbereich um eine weitere Population, indem sie erstmals die epistemologischen Überzeugungen von Frauen zu Wissen und Wissensprozess erforschte, darüber hinaus wurde ein weiterer wichtiger Bereich der Überzeugungen in die Entwicklungsmodelle integriert, die Überzeugungen zur Quelle des Wissens. Dieser Bereich betrifft Ansichten darüber, ob Wissen nur außerhalb der eigenen Person entsteht und von extern erhalten werden muss, ob Wissen ausschließlich internal ist, oder ob Wissen in Interaktion mit der Umwelt konstruiert wird (Hofer & Bendixen, 2012). Belenky et al. (1986) unterscheiden in ihrem *Women's Ways of Knowing* zwischen den folgenden fünf epistemologischen Perspektiven: *silence*, *received knowledge*, *subjective knowledge*, *procedural knowledge* (mit den Ausprägungen: *connected knowing* und *separate knowing*) und *constructed knowledge* (Hofer & Pintrich, 1997, S. 95).

Während Perrys Stichprobe ausschließlich aus männlichen und Belenkys ausschließlich aus weiblichen Studierenden bestand, entstand das *Epistemological Reflection Model* von Baxter Magolda und Porterfield (1985) und Baxter Magolda (1992) auf Basis von Interviews mit einer ausgewogenen Stichprobe von Männern und Frauen (Hofer & Bendixen, 2012). Auch hierbei handelt es sich um ein Entwicklungsstufenmodell. Baxter Magolda erfasste im Rahmen einer längsschnittlichen Interviewstudie über 5 Jahre, mittels des *Measure of Epistemological Reflection* (MER) (Baxter Magolda, 1987; Baxter Magolda & Porterfield, 1985), einerseits

die epistemologische Entwicklung und andererseits, wie epistemologische Annahmen die Interpretationen von Lernerfahrungen beeinflussen (Hofer & Pintrich, 1997, S. 97). Baxter Magolda und Kollegen unterscheiden in ihrem Modell zwischen *absolute knowing*, *transitional knowing*, *independent knowing* und *contextual knowing* (Baxter Magolda, 1992, 2004; Baxter Magolda, 2002; Baxter Magolda & Porterfield, 1985).

Die drei vorher genannten Entwicklungsmodelle (Baxter Magolda & Porterfield, 1985; Belenky et al., 1986; Perry, 1970) basieren auf phänomenologischen Interviews (Hofer & Bendixen, 2012). Zwei weitere Modelle, das *Reflective Judgment Model* (King & Kitchener, 1994; King, Kitchener, Davison, Parker & Wood, 1983) und das *Argumentative Reasoning* (Kuhn, 1991), die ebenfalls den Entwicklungsstufenmodellen zuzuordnen sind, arbeiten mit anderen qualitativen Interviewformen. Diese sind auf die Befragung zu Schlussfolgerungsprozessen bezogen.

King und Kitcheners Forschung (King & Kitchener, 1994, 2002, 2004; King et al., 1983) basiert einerseits auf Perry (1970) und andererseits auf den Arbeiten von Dewey (1933, 1938) zum reflektierten Schlussfolgern (reflective thinking). Sie untersuchten in ihren Interviewstudien, wie Individuen aus der Mittelschule bis zum mittleren Erwachsenenalter den Prozess des Wissens verstehen und wie sie dementsprechend ihre Überzeugungen zu schlecht strukturierten Problemen begründen (King & Kitchener, 1994, S. 13).

Kuhn (1991) legte ihren Fokus auf Denkprozesse bezüglich schlecht strukturierter Probleme im Alltag. Ihre Forschung zielte eigentlich auf argumentative Schlussfolgerungsprozesse, allerdings führte das einerseits indirekt zu Erkenntnissen über die individuellen Überzeugungen zu Wissen, andererseits befasste sich ein Teil der Studie auch explizit mit epistemologischen Perspektiven. Im Gegensatz zu den zeitlich früheren Forschungsbeiträgen wählte Kuhn (1991) eine viel breiter gestreute Stichprobe, sowohl aus dem akademischen Bereich als auch außerhalb. Ihr Forschungsdesign umfasst je vier bezüglich Geschlecht und Bildungsniveau (College, kein College) balancierten Altersgruppen (Teenager, 20er, 40er, 60er) mit je ca. 40 Probanden.

Die 5 oben genannten Entwicklungsstufenmodelle kennzeichnen die erste Phase der Forschung zu epistemologischen Überzeugungen. In dieser entwicklungspsychologischen Sichtweise werden epistemologische Überzeugungen als hoch integrierte, mehrdimensionale Konstrukte entworfen, die in Reaktion auf Bedingungen der Umwelt und Bildung entstehen (Hofer, 2016, S. 23 f.). Hofer und Pintrich (1997) konnten zeigen, dass die fünf Modelle sich in Bezug auf die angenommenen Stufen sehr ähnlich sind, auch wenn sie nicht alle die gleiche Anzahl an Stufen aufweisen. Auch der Entwicklungsprozess über die Stufen hinweg weist starke Ähnlichkeiten auf (ebd.). Generell haben diese Modelle gemeinsam, dass sie beschreiben, wie Individuen sich von einer dualistischen, absoluteren Sichweise auf Wissen und den Wissenserwerb zu einer subjektiveren, relativistischeren Position hin entwickeln, bevor sie eine evaluativere Perspektive erreichen können, die die Koordination der beiden vorherigen Perspektiven voraussetzt (Hofer, 2016, S. 23). Individuen mit absolutem bzw. dualistischem epistemologischen Denken betrachten Wissen als entweder eindeutig wahr oder falsch, von Autoritäten vermittelt, unveränderlich, aus Fakten bestehend und nicht zu hinterfragen. Auf relativistischer bzw. multiplizistischer Stufe sind Individuen der

Überzeugung, dass Wissen nichts anderes als Meinungen sind, die von jedem frei gewählt werden, veränderlich und nicht zu hinterfragen sind. Personen mit evaluativer Perspektive betrachten Wissen dagegen als sich entwickelnde, untereinander verbundene, komplexe Konzepte, die vom Individuum in Interaktion mit anderen konstruiert werden und die kritisch hinterfragt und abgewägt werden müssen (ebd.).

Alle 5 Modelle beschreiben die epistemologische Entwicklung als einen kognitiven Entwicklungsprozess, in dem sich das epistemologische Denken auf jeder Stufe qualitativ unterscheidet. Eine quantitative Ausprägung oder Veränderung in kognitiven Prozessen ist über diese Entwicklungsmodelle jedoch nicht beschreibbar. Die Stufen der Modelle sind sequenziell, das heißt, sie müssen der Reihe nach durchlaufen werden und der Wechel zur nächsten Stufe kann erst stattfinden, wenn die erste Stufe vollständig durchlaufen ist. Stufen zu überspringen ist nicht möglich. Die Entwicklung muss sich allerdings nicht über alle Stufen hinweg vollziehen, sondern kann bei einer Stufe enden. Den Entwicklungsstufenmodellen gemeinsam ist auch die Verwendung qualitativer Auswertungsmethoden. In ihren Erhebungsmethoden unterscheiden sich die Modelle jedoch (ebd.).

Auffällig ist, dass die Entwicklungsstufenmodelle in der Vergangenheit häufig als eindimensional bezeichnet wurden (z.B. Müller, 2009). Das ist allerdings nicht der Fall. In den Modellen von Baxter Magolda (1992) und King und Kitchener (1994) wird sogar explizit Bezug auf verschiedene Dimensionen genommen, während diese bei Kuhn (1991) und Perry (1970) nicht explizit enthalten sind, aber von Hofer und Pintrich (1997) abgeleitet werden konnten (Hofer & Pintrich, 1997, S. 118). Als mehrdimensional können diese Konzeptionen insofern bezeichnet werden, als die einzelnen Stufen (z.B. Absolutismus zu Relativismus) sich in Bezug auf mehrere Aspekte epistemologischer Überzeugungen qualitativ unterscheiden. Die Entwicklung von einer Stufe in die nächste geht einher mit synchronen Veränderungen in verschiedenen Dimensionen, wie z.B. Quelle, Sicherheit und Rechtfertigung des Wissens (Hofer & Bendixen, 2012, S. 239). Der Übergang von der Stufe *Absolutism* zu *Relativism* ist also nicht nur durch eine Veränderung in Bezug auf die Überzeugungen zur Quelle des Wissens gekennzeichnet, sondern gleichzeitig auch durch veränderte Überzeugungen zur Wissensbegründung. Jede Stufe integriert also inhaltlich verschiedene Facetten der epistemologischen Überzeugungen, die allerdings nicht getrennt voneinander betrachtet werden können, da sich die Veränderung synchron vollzieht. Die Erhebung bei diesen Modellen erfolgt also notwendigerweise immer eindimensional (Hofer, 2016; Hofer & Pintrich, 1997).

Der zentrale Beitrag der ersten Forschungsphase liegt darin, dass hier die Basis für die epistemologische Forschung gelegt wurde, indem epistemologische Überzeugungen als psychologisches Konstrukt erstmals beschrieben wurden (Hofer, 2016, S. 23 f.). Dabei handelt es sich um ein Konstrukt, dessen Bedeutung für die höhere Bildung und Verbindung zu kritischem Denken und Argumentieren bereits in der ersten Forschungsphase deutlich wurde. Die Forschenden wendeten hier offene Erhebungsmethoden an, die es ermöglichten, eine phänomenologische Sicht auf die Bedeutungskonstruktion von Lernern zu bekommen (ebd.). Es gelang ihnen, Fragen zu konstruieren, die die epistemologische Perspektive von

Lernern bezüglich komplexen Problemen erfassbar machen. Die entwickelten Modelle stellen nützliche Heuristiken für das Verstehen der epistemologischen Entwicklung während des Colleges dar. Sie versetzen andere Forschende in die Lage, diese Erkenntnisse in die Praxis der Hochschulbildung zu übertragen und so den Wert fortgeschrittenen epistemologischen Denkens zu erkennen und zu erforschen, wie die akademische Bildung diese fördern könnte (ebd.). Diese Entwicklungsstufenmodelle stellen bis heute die Basis eines eigenen Forschungsstrangs dar (z.B. Baxter Magolda, Creamer & Meszaros, 2010; King & Baxter Magolda, 2005; Kuhn, Iordanou, Pease & Wirkala, 2008; Kuhn & Pearsall, 2000; Kuhn, Zillmer, Crowell & Zavala, 2013; Tabak & Weinstock, 2008).

### 2.1.2 Zweite Forschungsphase – Explizite Mehrdimensionalität

Schommer (1990) leitete mit ihrem *Epistemological Belief System* eine neue Phase in der Forschung zu epistemologischen Überzeugungen ein. Sie betrachtete die epistemologischen Überzeugungen erstmals nicht nur als mehrdimensionales Konstrukt, sondern als ein System epistemologischer Überzeugungen mit mehr oder weniger unabhängigen Dimensionen. Schommer (1990) nahm a priori fünf kontinuierliche Dimensionen an: *Stability of Knowledge*, *Source of Knowledge*, *Structure of Knowledge*, *Ability of Learning* und *Speed of Learning*. Diese bilden unterschiedliche Teilaspekte des Konstrukts ab und entwickeln sich in ihrer Ausprägung mehr oder weniger unabhängig voneinander. Individuen befinden sich also nach Schommers Sichtweise nicht mehr zu einem Zeitpunkt in einer bestimmten Stufe epistemologischer Überzeugungen, wie dies bei den Entwicklungsstufenmodellen der Fall war, sondern verfügen über verschiedene Ausprägungen epistemologischer Überzeugungen in jeder der fünf verschiedenen Dimensionen. Das bedeutet, dass in Bezug auf die individuellen epistemologischen Überzeugungen nun sowohl qualitative Unterschiede als auch quantitative Unterschiede in der Merkmalsausprägung beschreibbar wurden. Damit einher geht auch die quantitative Erhebung mittels Selbstauskünften über likertskalierte Items, die im Bereich epistemologischer Überzeugungen auf M. P. Ryan (1984) zurückgeht (Schommer-Aikins, 2004, S. 20 f.). Hier steht nicht mehr die Ontogenese epistemologischer Überzeugungen im Vordergrund des Forschungsinteresses, sondern vielmehr die interne Struktur epistemologischer Überzeugungen als System von Überzeugungen. Ihre Konzeption stellt daher einen persönlichkeitspsychologischen Ansatz dar (Hofer, 2001). In den Fokus des Forschungsinteresses rückten durch die ökonomischere Erhebung, die große Stichproben ermöglichte, nun auch die Beziehungen der individuellen Überzeugungen mit anderen Konstrukten der pädagogischen Psychologie. Schommer (1990) betrachtet spezielle Überzeugungen zum individuellen Lernverhalten, im Rahmen der Dimensionen *Ability of Learning* und *Speed of Learning*, sogar als Teil des Konstrukts epistemologischer Überzeugungen. Faktorenanalytische Untersuchungen (Schommer, 1990, 1993a; Schommer et al., 1992) der fünf Dimensionen führten zu vier Faktoren, die die folgenden 4 Dimensionen abbilden: *Fixed Ability*, *Speed of Learning*, *Simple Knowledge*, *Certainty of Knowledge* (Schommer-Aikins, 2004). Schommer konnte zeigen, dass die epistemologischen Überzeugungen Verbindungen mit Verstehen (Schommer, 1990), Leistung und Lernstrategienutzung (Schommer et al., 1992) aufweisen.

Diese Forschungsphase ist auch dadurch gekennzeichnet, dass Schommers Erhebungsinstrument (Schommer, 1990; Schommer-Aikins & Hutter, 2002) zahlreich angewendet wurde (z.B. Bendixen, Dunkle & Schraw, 1994; Schommer-Aikins & Hutter, 2002; Schraw et al., 2002) und viele gekürzte (z.B. Kardash & Noel, 2000; Kardash & Scholes, 1996; Mori, 1999; Qian & Alvermann, 1995), modifizierte und übersetzte (Chan & Elliott, 2000, 2002, chinesisch; Clarebout, Elen, Luyten & Bamps, 2001, niederländisch) Fragebogen-Versionen mit unterschiedlichen Konfigurationen bezüglich der Dimensionen entwickelt und getestet wurden. Auf Basis von Schommers Konzeption bzw. unter Rückgriff auf ihre Items entstanden in dieser Zeit auch neue Instrumente, wie *Jehng's Epistemological Questionnaire* (JEQ, Jehng et al., 1993) und das *Epistemological Belief Inventory* (EBI, Schraw, Dunkle & Bendixen, 1995).

Hofer und Pintrich (1997) leisteten im Anschluss einen bedeutenden Beitrag, indem sie die verschiedenen Forschungsstränge zu epistemologischen Überzeugungen im Rahmen ihres Reviews strukturierten und auf Basis der Ergebnisse eine eigene Konzeption, die *Epistemological Theories*, vorschlugen. Sie analysierten alle bedeutenden Konzeptionen epistemologischer Überzeugungen (Baxter Magolda, 1992; Belenky et al., 1986; King & Kitchener, 1994; Kuhn, 1991; Perry, 1970; Schommer, 1990), die bis dahin veröffentlicht worden waren, identifizierten die enthaltenen Dimensionen, inhaltliche Gemeinsamkeiten und Unterschiede in Bezug auf die Dimensionen und setzten die einzelnen Stufen, Positionen bzw. Ebenen der fünf Entwicklungsstufenmodelle zueinander in Relation. Sie schlugen im Rahmen ihres Reviews zwei Überkategorien mit insgesamt vier Kerndimensionen epistemologischer Überzeugungen vor, die sie als Schnittmenge zwischen den verschiedenen Forschungskonzeptionen identifiziert hatten. Bei den Überkategorien handelt es sich um Überzeugungen zur *Natur des Wissens* (*nature of knowlegde*) und zur *Natur des Wissensprozesses* (*nature of knowing*) (Hofer & Pintrich, 1997). Diese beiden zentralen Aspekte werden bis heute als Kern des Konstrukts epistemologischer Überzeugungen gesehen und für die Definition des Konstrukts verwendet. Als Kerndimensionen *epistemologischer Theorien* – wie sie von Hofer und Pintrich (1997) bezeichnet werden – identifizierten sie zum einen die beiden Dimensionen *Sicherheit des Wissens* und *Einfachheit des Wissens*, die der Kategorie *nature of knowledge* zugeordnet werden und zum anderen die beiden Dimensionen *Quelle des Wissens* und *Begründung des Wissens*, die der Kategorie *nature of knowing* angehören (Brownlee, Ferguson & Ryan, 2017; Hofer & Pintrich, 1997). Auch diese vier zentralen Dimensionen sind bis heute als Kerndimensionen anerkannt (siehe z.B. neuere Modelle: *Theory of Intregrated Domains in Epistemology (TIDE) framework*, Muis et al., 2006; *Integrative Personal Epistemology Model*, Bendixen & Rule, 2004; *Embedded Systemic Model*, Schommer-Aikins, 2004).

Im Gegensatz zu Schommer (1990), die epistemologische Überzeugungen als System von Eigenschaften und als mehr oder weniger unabhängige Dimensionen beschreibt, gehen Hofer und Kollegen davon aus, dass die individuellen Überzeugungen zu Wissen und zum Wissensprozess eher in Form von persönlichen Theorien organisiert sind (Hofer, 2001, S. 360 f.). Diese Sichtweise übernimmt die explizite Multidimensionalität, wie Schommer sie

vorschlägt, geht aber von einer höheren Integration zwischen den individuellen Überzeugungsbereichen aus. Die entspricht auf eine gewisse Art wieder mehr den Entwicklungsstufenmodellen, die ja innerhalb einer jeden Stufe kohärent verschiedene Dimensionen als integriert darstellten (Hofer, 2001, S. 361).

Hofers Konzeption steht im Einklang mit der Annahme, dass die epistemologischen Überzeugungen einer Person sich in Abhängigkeit der Disziplin (z.B. Psychologie vs. Physik) unterscheiden können (Hofer, 2001; Hofer & Pintrich, 1997). Diese ersten Vorstöße in die Erforschung intraindividueller, domänenspezifischer epistemologischer Überzeugungen (Donald, 1995; Hofer, 2000) stellen in dieser Zeit eine wichtige neue Forschungsentwicklung dar. Hofer (2000) entwickelte auf Basis ihrer Konzeption (Hofer & Pintrich, 1997) den *Discipline-Focused Epistemological Belief Questionnaire* (DEBQ), der es ermöglichen soll, die epistemologischen Überzeugung in Abhängigkeit von der spezifischen Disziplin zu erheben. Etwas später gehen sie davon aus, dass individuelle epistemologische Theorien vom jeweiligen Bereich abhängig sind und dieselbe Person sowohl über allgemeine und gleichzeitig auch über andere domänenspezifische epistemologische Überzeugungen verfügen kann (Hofer, 2001, S. 362). Auch auf Basis der Konzeption von Hofer und Pintrich (1997) entstanden verschiedene Fragebögen, die allerdings zeitlich in der dritten Forschungsphase zu verorten sind (*Internet-Specific Epistemological Questionnaire*, ISEQ, Bråten & Olaussen, 2005; Epistemology Survey, Karabenick, 2005).

Zusammenfassend betrifft eine der zentralen Fragen dieser Forschungsphase nach wie vor die Konzeptualisierung des Konstrukts epistemologischer Überzeugungen in Theorie und Forschung. Darüber hinaus liegt der Fokus des Forschungsinteresses nun auch auf der Mehrdimensionalität, der Unterscheidung zwischen domänenspezifischen und domänenübergreifenden Überzeugungen sowie im Besonderen der Rolle epistemologischer Überzeugungen beim Lehren und Lernen. Die Bedeutung epistemologischer Überzeugungen im Lehr-Lernprozess betrifft dabei sowohl die Lehrerbildung, praktizierende Lehrkräfte sowie Studierende im Allgemeinen und Schülerinnen und Schüler (Brownlee, Ferguson & Ryan, 2017; Hofer, 2016). Dies geht sehr deutlich aus einigen zentralen Werken dieser Zeit hervor wie den Sammelbänden von Hofer und Pintrich (2002) und Sinatra und Pintrich (2003).

### 2.1.3 Dritte Forschungsphase – Ausdifferenzierung

Die Forschung zu epistemologischen Überzeugungen nahm in der dritten Forschungsphase stark zu und erstreckt sich nun über die verschiedensten akademischen Kontexte und Fachbereiche wie die Pädagogische Psychologie, Entwicklungspsychologie, Philosophie, Lehrerbildung, Fachdidaktiken, Hochschulbildung, Bildungsforschung und dringt in die unterschiedlichsten Kulturen sowie Altersgruppen und Berufsgruppen vor (Hofer, 2016, S. 20). Während es Hofer und Pintrich 1997 noch gelang den Forschungsstand in einem einzelnen Review umfassend abzubilden (Hofer, 2016, S. 20), existieren nun unzählige Artikel, Sonderhefte, Kapitel und Sammelbände. Einige bedeutende Werke seien hier genannt: Das Review von Hofer (2008) und der Sammelband von Khine (2008) befassen sich beispielsweise mit Beiträgen zu epistemologischen Überzeugungen über verschiedene Kulturen, während das

Review von Chinn et al. (2011) sowie das Werk von Moshman (2015) die wissensbezogenen Überzeugungen mit ihren Wurzeln in der Philosophie verknüpfen. Muis et al. (2006) befassen sich in ihrem Review mit der Disziplinspezifität von epistemologischen Überzeugungen anhand von 19 Studien und schlagen auf Basis der Forschungsergebnisse ein Modell vor, dass die verschiedenen Ansätze des Forschungsbereichs integriert. Im Sammelband von Khine (2008) sind ebenfalls Beiträge zur Domänenspezifität zu finden. Das Sonderheft des *Educational Psychologist* (2017, Heft 4, Jahrgang 52) sowie die Sammelwerke von Bendixen und Feucht (2010), Brownlee et al. (2011), Fives und Gill (2014) und Schraw et al. (2017) beleuchten epistemologische Überzeugungen in Bezug auf die Lehrerbildung, Lehrkräfte und den Kontext Schule. Der Sammelband von Elen et al. (2011) enthält Beiträge verschiedenster Art zur *cognitive flexibility*. Die Beiträge im Handbuch von Greene et al. (2016) beleuchten verschiedene psychologische Perspektiven sowie die Perspektiven verschiedener Fachdisziplinen auf epistemologische Überzeugungen (*epistemic cognition*) und betreffen auch die epistemologischen Überzeugungen von Lehrkräften, Interventionen zur Veränderung derselben sowie Aspekte der Messung und Modellierung.

Wie im Rahmen der ersten Forschungsphase bereits erwähnt, wird die Forschung zur Entwicklung epistemologischer Überzeugungen über die Stufen Absolutismus, Relativimus bis hin zu Evaluatismus über die Forschungsphasen hinweg weitergeführt. Nach den Erkenntnissen neuerer Beiträge in diesem Forschungsstrang wird nun davon ausgegangen, dass die Entwicklung nicht mehr nur in eine Richtung möglich ist, sondern auch eine Regression in frühere Entwicklungsstufen auftreten kann, hervorgerufen durch z.B. Wechsel zwischen Bildungskontexten, wie von der Schule zur Universität (z.B. Bendixen, 2002; Bendixen & Rule, 2004; Chandler, Hallett & Sokol, 2002).

Die epistemologischen Überzeugungen wurden in der ersten Forschungsphase hauptsächlich qualitativ über Interviews erfasst, während sich in der zweiten Forschungsphase Likertskalen zur Erhebung etablierten. In dieser dritten Forschungsphase wurden auch aufgrund von Kritik an den bisherigen Erhebungsverfahren, viele alternative Erhebungsmethoden eingesetzt. Stahl und Bromme (2007) setzen mit dem *Connotative Aspects of Epistemological Beliefs* (CAEB) eine ganz andere Erhebungsmethode ein. Sie erfassen die epistemologischen Überzeugungen von Studierenden bezüglich verschiedener Bereiche der Naturwissenschaft mittels seminatischem Differential. Muis, Trevors und Chevrier (2016) kombinieren hingegen quantitative und qualitative Methoden, um das TIDE Modell zu prüfen. Auch Bromme (2010) nutzt Mixed Methods zur Erhebung. In anderen Studien wurden dagegen Laut-Denken-Protokolle eingesetzt (Ferguson, Bråten & Strømsø, 2012; Hofer, 2004a; Mason et al., 2010b) sowie Unterrichtsbeobachtungsmethoden (Elby & Hammer, 2010; Hofer, 2004b), Card-Sorting Methoden und Netzwerkanalysen (Peters-Burton & Baynard, 2013). Mason (2016) und Kelly (2016) bieten Reviews zu den neueren Erhebungsmethoden.

In den Disziplinen Mathematik (vgl. Depaepe, de Corte & Verschaffel, 2016), Naturwissenschaften (vgl. Elby, Macrander & Hammer, 2016) und Geschichte (vgl. VanSledright & Maggioni, 2016) entstanden über die Zeit eigene Forschungsstränge, die sich mit epistemologischen Überzeugungen bezüglich der spezifischen Disziplin befassen und in der dritten

Forschungsphase fortgeführt werden. Diese sind im Bereich der Fachdidaktiken angesiedelt. Auch zum Bereich *Literary Reasoning* (vgl. Lee, Goldman, Levine & Magliano, 2016) entwickelte sich ein eigener Forschungsstrang, der bis heute aktiv ist und die Rolle der individuellen epistemologischen Überzeugungen beim Umgang, Lesen, Verstehen und Schlussfolgern aus Texten untersucht.

Im Rahmen der dritten Forschungsphase rücken neben den Studierenden, die in den vorangegangenen Forschungsphasen häufig die Probanden darstellten, nun auch andere Altersgruppen in den Fokus der Forschenden. Studien zu den epistemologischen Überzeugungen von Vorschulkindern (z.B. Barth, Bhandari, Garcia, MacDonald & Chase, 2014; Burr & Hofer, 2002; Wildenger et al., 2010) und Grundschulkindern (z.B. Conley et al., 2004; Elder, 2002) erweitern das Feld (vgl. Hofer, 2016). Auch unterschiedliche Berufsgruppen, wie beispielsweise Erzieherinnen und Erzieher (z.B. Plöger-Werner, 2015) und Auszubildende (z.B. Mokwinski, 2011; Zinn, 2013) sowie das berufliche Lernen an sich (vgl. Bernholt et al., 2017) und ganz zentral Lehrkräfte (vgl. Bendixen & Feucht, 2010; Brownlee et al., 2011; Brownlee, Schraw & Berthelsen, 2013; Fives & Gill, 2014; Schraw et al., 2017) sind nun von Interesse (vgl. Hofer, 2016).

Die Forschung zu epistemologischen Überzeugungen und deren Verbindungen mit dem Lehr- und Lernprozess weitet sich über die USA hinaus nach Europa, Asien und in viele Regionen der Welt, wie Ägypten, Saudi Arabien, Brasilien, Argentinien etc. aus (Hofer, 2008). Gleichzeitig wird auch vermehrt der kulturelle Einfluss auf die Entwicklung epistemologischer Überzeugungen erforscht (Hofer & Bendixen, 2012, S. 235), da dieser eine bedeutende Rolle spielt (z.B. Gottlieb, 2007; Haerle & Bendixen, 2008; Qian & Pan, 2002). Hofer (2008) stellt in ihrem Review über die verschiedensten Kulturen hinweg Forschungsbeispiele und Implikationen für das Konstrukt epistemologischer Überzeugungen vor. Der Sammelband von Khine (2008) enthält einige Beiträge zu empirischen Studien über kulturspezifische epistemologische Überzeugungen. Die Kultur als eigens zu berücksichtigendes System findet sich in Schommers weiterentwickeltem Modell, dem *Embedded Systemic Model* Schommer-Aikins (2004, S. 23 ff.). Es bettet die epistemologischen Überzeugungen in einen größeren kulturellen Kontext ein.

Insgesamt werden die Forschungsstränge zur Kontextabhängigkeit, Domänenspezifität und Mehrdimensionalität in dieser dritten Forschungsphase weitergeführt und stark ausgeweitet (Hofer, 2016). Die Existenz domänenspezifischer epistemologischer Überzeugungen, d.h. unterschiedlicher Überzeugungen in Abhängigkeit von der akademischen Disziplin, wird mittlerweile aufgrund der vielen Forschungsergebnisse im Forschungsbereich allgemein anerkannt (Buehl & Alexander, 2005; Hofer, 2006a; Muis et al., 2006). Nach der Einschätzung Hofers (2006a), schlossen Muis et al. (2006) diese Debatte zur Domänenspezifität mittels ihres viel rezipierten Reviews über 19 empirische Studien ab. Gleichzeitig wird auch eine mögliche Themenspezifität epistemologischer Überzeugungen näher untersucht (z.B. Bråten, Strømsø & Samuelstuen, 2008).

Die Kontextspezifität epistemologischer Überzeugungen wird auch in der theoretischen Entwicklung des Konstrukts, beispielsweise im Rahmen des *Epistemological Resources* Modell (Hammer & Elby, 2002; Louca et al., 2004) aufgegriffen. Hammer und Elby (2002) und

Louca et al. (2004) betrachten epistemologische Überzeugungen nicht als System von Eigenschaften oder persönlichen Theorien, sondern als kognitive Struktur von epistemologischen Ressourcen (*epistemological resources*), die dazu dienen Wissen zu verstehen. Sie stimmen dagegen mit Schommer (1990) und Hofer und Pintrich (1997) darin überein, dass die explizite Annahme verschiedener Dimensionen gegenüber den Entwicklungsstufenmodellen das Konstrukt adäquater beschreiben kann (Hammer & Elby, 2002, S. 171). Nach dieser Konzeption aktivieren unterschiedliche Kontexte (z.B. Lernkontexte) unterschiedliche epistemologische Ressourcen. Diese sind demnach hoch kontextabhängig und disziplinspezifisch (Hammer & Elby, 2002; Louca et al., 2004).

Buehl und Kollegen konzipieren epistemologische Überzeugungen ebenfalls als disziplinspezifisch (Buehl & Alexander, 2001, 2006; Buehl et al., 2002). Sie bezeichnen epistemologische Überzeugungen insofern als „multilayered" (Buehl & Alexander, 2001, S. 385), als dass Individuen zunächst über übergreifende Überzeugungen zum Wissen verfügen, aber sich aus diesen allgemeinen epistemologischen Überzeugungen durch die Auseinandersetzung mit spezifischen akademischen Disziplinen disziplinspezifische Überzeugungen zu den betreffenden akademischen Wissensbereichen entwickeln. Dementsprechend wird diese Konzeption als *Beliefs about Academic Knowledge* bezeichnet (Buehl & Alexander, 2001, 2006; Buehl et al., 2002).

Mit dem *Theory of Integrated Domains in Epistemology Framework* (TIDE) gelingt es Muis et al. (2006), verschiedene prominente Perspektiven zu epistemologischen Überzeugungen sowie die Annahme, dass epistemologische Überzeugungen sowohl domänenübergreifend als auch domänenspezifisch sind, zu vereinen. Gleichzeitig werden die domänenübergreifenden und -spezifischen Überzeugungen im übergreifenden soziokulturellen Kontext sowie im akademischen und instruktionalen Kontext verortet. Muis et al. (2006) integrieren unter anderem auch die Annahmen zur zeitlichen Entwicklung epistemologischer Überzeugungen von Absolutismus, Relativimus bis hin zu Evaluatismus. Allerdings gehen sie nicht von einer linearen Entwicklung in eine Richtung, sondern von einer spiralförmigen Entwicklung aus, die Rückentwicklungen im Sinne von Chandler et al. (2002), Bendixen (2002) und Bendixen und Rule (2004) erlaubt. Sie vermuten, dass solche Rückentwicklungen durch starke Veränderungen im Bildungskontext hervorgerufen werden könnten, weil sich die individuellen Erwartungen ändern und sowohl die Inhalte als auch die Bildungsumgebung komplexer werden (Muis et al., 2006).

Ein weiteres Modell, das nicht nur Perspektiven der expliziten Mehrdimensionalität epistemologischer Überzeugungen (z.B. im Sinne von Hofer, 2000) mit den Entwicklungsmodellen (z.B. King & Kitchener, 1994, 2004) integriert sondern auch externale Einflussfaktoren berücksichtigt, ist das *Integrated Modell of Personal Epistemology* von Bendixen und Rule (2004). Das Modell beschreibt ganz zentral den Mechanismus der Veränderung epistemologischer Überzeugungen und welche Elemente bei der Entwicklung wie zusammenwirken (Bendixen & Rule, 2004).

Wie sich beispielsweise am Modell von Bendixen und Rule (2004) zeigen lässt, entwickelt sich auch die Verknüpfung der epistemologischen Überzeugungen mit dem Konstrukt der Kognition und der Metakognition in dieser Forschungsphase deutlich weiter (Hofer,

2016). Zwar begannen schon Royce und Kollegen (Royce, 1978; Royce & Mos, 1980; Royce & Smith, 1964), Kitchener (1983) (erste Forschungsphase) sowie Kuhn und Kollegen (Kuhn, 1999a, 1999b, 2000; Kuhn, Cheney & Weinstock, 2000) (zweite Forschungsphase) diese Beziehung genauer zu untersuchen, doch nun wird sie von verschiedenen Forscherinnen und Forschern aufgegriffen und deutlich ausdifferenziert (z.B. von Barzilai & Zohar, 2014, 2016; Bendixen & Rule, 2004; Bromme, 2010; Hofer, 2004a, 2005; Mason et al., 2010b).

Weitere Beziehungen zu anderen Konstrukten betreffen u.a. das selbstgesteuerte Lernen (z.B. Bromme, 2010; Pieschl et al., 2013), *Conceptual Change* (z.B. Mason et al., 2010b; Mason & Boscolo, 2004), Motivation (z.B. Buehl & Alexander, 2005), Volition (z.B. Rule & Bendixen, 2010), *need for cognition* und *need for closure* (z.B. DeBacker & Crowson, 2006) und das Argumentieren (z.B. Kuhn et al., 2013; Mason, Scirica & Salvi, 2006; Yang & Tsai, 2010). Hofer und Bendixen (2012) bieten einen guten Überblick zu den Beziehungen zwischen epistemologischen Überzeugungen zu anderen Konstrukten.

Gleichzeitig besinnen sich die Forscherinnen und Forscher dieses Feldes auf die philosophischen Ursprünge der epistemologischen Überzeugungen (z.B. Moshman, 2015; Muis et al., 2006) und es entstehen Modelle, die epistemologische Überzeugungen mit der philosophischen Grundlage der Epistemologie verbinden (Hofer, 2016, S. 25). Dies führt zu einer Ausweitung des Konstrukts in Bezug auf verschiedene Aspekte der Wissensbegründung. Hier sei auf die Ansätze von Chinn et al. (2011), Greene et al. (2008) und Murphy (2003) verwiesen. Im Rahmen dieser neuen Modelle wird nun teilweise der Begriff *epistemic cognition* für epistemologische Überzeugungen verwendet. Ein Beispiel hierfür ist das *Aims, Ideals and Reliable processes* (AIR) Modell (Chinn & Rinehart, 2016; Chinn, Rinehart & Buckland, 2014), das auf ihrem Review (Chinn et al., 2011) basiert.

Aus den Verbindungen der epistemologischen Überzeugungen zur Philosopie entsteht die Forschung zu sozialen Dimensionen des Wissens (Hofer, 2016). Dazu gehört die Forschung zum Umgang mit Meinungen und Sichtweisen von anderen Personen, Gruppenentscheidungen und den Einflüssen auf das gesellschaftliche Wissen, wie z.B. das öffentliche Verständnis von Naturwissenschaften (z.B. Sinatra, Kienhues & Hofer, 2014) sowie institutioneller Epistemologie und epistemologischen Überzeugungen (z.B. Goldman & Blanchard, 2012; Sinatra et al., 2014).

In dieser Forschungsphase entstehen darüber hinaus zahlreiche Interventionen zur Veränderung epistemologischer Überzeugungen von Schülerinnen und Schülern, Studierenden und spezifisch Lehramtsstudierenden sowie Lehrkräften (siehe hierzu Bråten, 2016; Brownlee et al., 2016; Muis et al., 2016; Parkinson & Maggioni, 2017). Das 3R-EC Framework (*Reflection, Reflexivity and Resolved Action for Epistemic Cognition*, Brownlee et al., 2016) beispielsweise befasst sich speziell mit der Veränderung der epistemischen Kognition von Lehrkräften in kontextspezifischen Lehr-Lernkontexten (Brownlee, Ferguson & Ryan, 2017). Es basiert auf M. Ryan und Bourke (2013) und integriert das AIR Modell (Chinn & Rinehart, 2016; Chinn et al., 2014). Brownlee, Ferguson und Ryan (2017, S.9) schlagen vor, dass dieses Modell und die damit einhergehende Forschung zur systematischen Veränderung der epistemischen Kognition von Lehrkräften vielleicht gar den Anfang einer vierten Forschungsphase einleitet.

Als eine der zentralen Entwicklungen dieser Forschungsphase hebt Hofer (2016, S. 27) die Forschung bezüglich des Einflusses der epistemologischen Überzeugungen von angehenden und praktizierenden Lehrkräften auf den Lernprozess und die Überzeugungen der Schülerinnen und Schüler hervor (z.B. Bendixen & Feucht, 2010; Brownlee et al., 2011, 2013; Fives & Gill, 2014; Schraw et al., 2017).

## 2.2 Modelle epistemologischer Überzeugungen

In der Darstellung der Forschungsentwicklung wurde bereits auf verschiedene Modelle epistemologischer Überzeugungen verwiesen. All diese Konzeptualisierungen stellen Antworten auf die Fragen dar, was epistemologische Überzeugungen sind, wie sie definiert werden und wo Grenzen zu anderen Konstrukten gezogen werden, d.h. was nicht mehr zum Konstrukt epistemologischer Überzeugungen gehört. Im Folgenden werden die im Rahmen dieser Instrumententwicklung theoretisch relevanten Modelle vorgestellt. Für einen umfassenderen Überblick über die früheren Modelle sei auf den viel rezipierten Beitrag von Hofer und Pintrich (1997) und den synoptischen Überblick von Zinn (2013, S. 61 ff.) verwiesen.

Zentral für diese Arbeit sind im besonderen die *Epistemological Theories* von Hofer und Pintrich (1997) sowie Hofer (2000; 2001). Da dieses theoretische Modell allerdings in seinen verschiedenen Aspekten auf früheren Modellen basiert, wird zunächst ein Überblick über die betreffenden zugrundeliegenden und inhaltlich verbundenen Modelle gegeben. Den Anfang bildet das *Scheme of intellectual and Ethical Development* von Perry (1970, 1981). Es stellt als erstes Entwicklungsstufenmodell eine wichtige Basis für die nachfolgende Forschung und Instrumententwicklung in diesem Forschungsbereich dar. Im Anschluss wird das *Reflective Judgment Model* von King und Kitchener (1994) sowie das Modell zum *Development of Epistemological Understanding* von Kuhn (1991) vorgestellt. Nach diesen drei zentralen Entwicklungsstufenmodellen wird in das erste pädagogisch-psychologische Modell, das *Epistemological Belief System* von Schommer (1990), eingeführt. Es legte die Basis für die folgenden pädagogisch-psychologischen Modelle und die explizite Annahme und empirische Erforschung verschiedener Überzeugungsdimensionen. Im Anschluss an diese vier Modelle wird schließlich die zentrale Konzeptualisierung von Hofer und Pintrich (1997) beschrieben, die *Epistemological Theories*.

### 2.2.1 Scheme of Intellectual and Ethical Development von Perry

Perrys längsschnittliche, qualitative Forschung (1970; 1981) zielte darauf ab, die unterschiedlichen Reaktionen und den verschiedenartigen Umgang der Studierenden mit dem pluralistischen, intellektuellen und sozialen Umfeld der Universität zu erklären. Er wollte die Ursache dafür identifizieren, dass verschiedene Studierende sich die Bedeutung ihrer Lernerfahrungen unterschiedlich erklären (Perry, 1970). Ursprünglich ging er davon aus, stabile Persönlichkeitsunterschiede zu finden, die die Ursache seien. Er entwickelte dazu eine Checkliste für Bildungswerte (CLEV, Checklist of Educational Values), die im Rahmen von Interviews mit männlichen Harvard-Studenten eingesetzt wurde. Auf Basis der resultierenden Ergebnisse des CLEV wählte er Studierende, die im Verlauf des Studienjahres starke Änderungen in ihren Ansichten gezeigt hatten oder die extreme Ausprägungen erreichten, für die weiteren Interviews aus. Daraufhin führte er mit den ausgewählten Probanden über einen Zeitraum von vier Studienjahren (in den späten 1950er und Anfang der 1960er Jahre) offene Interviews (ebd.). Die Auswertung der offenen Interviews führte schließlich zu der

Veröffentlichung des Stufenmodells (Perry, 1970), das die intellektuelle und ethische Entwicklung in der späten Adoleszenz bis hin ins Erwachsenenalter (17- bis 22-Jährigen) im Verlauf der akademischen Ausbildung beschreibt.

Auch wenn Perry selbst im Zusammenhang mit seinem Stufenmodell nicht von epistemologischer sondern von intellektueller und ethischer Entwicklung spricht, beschreibt sein Modell ein sich weiterentwickelndes Verständnis über das Wissen, den Wissensprozess und Lernen (Moore, 2002). Das *Perry Scheme* stellt nach Einschätzung Hofer und Pintrichs (1997) ein Modell dar „for interpreting students' epistemological responses to the college environment" (S. 93). Es wird deshalb im Forschungsbereich übereinstimmend als das erste epistemologische Modell bezeichnet und stellt die Grundlage für viele weitere Modelle dar.

Das *Perry Scheme* beschreibt neun Positionen intellektueller und ethischer Entwicklung, die den folgenden vier Kategorien zugeordnet werden: Dualismus (*dualism*), Multiplizismus (*multiplizism*), kontextueller Relativismus (*contextual relativism*) und Selbstverpflichtung im Relativismus (*commitment within relativism*) (Perry, 1970). Die erste und die letzte der 9 Positionen basieren nicht auf den Forschungsergebnissen, sondern stellen hypothetische Erweiterungen Perrys dar (Hofer & Pintrich, 1997). Die folgende Beschreibung basiert auf der Darstellung von Perry (1981, S. 76 ff.) und Moore (2002, S. 19 ff.).

**Dualismus (*Dualism*)**

Die Positionen des Dualismus zeichnen sich durch eine Dichotomisierung in den Ansichten aus. Es wird strikt zwischen richtig und falsch, gut und schlecht, wir und sie, Erfolg und Misserfolg, etc. getrennt. Für jede Frage und jedes Problem gibt es eine richtige Antwort bzw. Lösung. Experten verfügen über diese Antworten. Wissen wird als quantitativ gesehen, d.h. Lernen besteht darin, die richtigen Antworten und Lösungen mittels großer Anstrengung auswendig zu lernen. Die Fähigkeit zu handeln und zur Entscheidung wird external in Autoritäten, Testergebnissen und dem *richtigen* Beruf gesehen (Perry, 1981, S. 79).

**Position 1: *Basic duality***

Eine Person in dieser Position verfügt über eine vollständig unhinterfragte und unreflektierte Sichtweise auf die Wahrheit als *absolute Wahrheit*. Dementsprechend wird eine klare Grenze zwischen absolut wahr und falsch gesehen. Die Person identifiziert sich absolut und unhinterfragt mit Autoritätspersonen (Eltern, Lehrkräfte, Kirche). Alternative Sichweisen werden nicht toleriert (Moore, 2002, S. 20).

**Position 2: *Multiplicity prelegitimate***

Eine Person in dieser Position sieht zwar, dass es verschiedene Perspektiven und Ansichten gibt, die aber schlicht falsch sind. Das Denken ist charakterisiert durch Dichotomie und Dualismus, in dem Sinne *wir-richtig-gut* im Gegensatz zu *sie-falsch-schlecht* oder Abwandlungen davon. Die Welt wird im Sinne dieses Schubladendenkens in Richtiges und Falsches unterteilt. Das Richtige vom Falschen zu unterscheiden ist nach dieser Überzeugung im Allgemeinen einfach (Moore, 2002, S. 20).

**Multiplizismus (*Multiplicity*)**

Position 3 und 4 werden als *Multiplicity* bezeichnet, da in diesen Phasen die Konfrontation und der Umgang mit Diversität und dementsprechend Unsicherheit in den verschiedensten Bereichen die zentrale Rolle spielt (Moore, 2002, S. 20). Diversität in Meinungen und Werten wird nur in Bereichen als legitim anerkannt, in denen die *richtige* Lösung/Antwort bisher noch nicht bekannt ist. Meinungen sind unverbunden und ihnen liegt kein System zugrunde. Im Multiplizismus sind Personen noch nicht in der Lage Meinungen zu werten. Daraus folgt die Ansicht, dass jede Meinung gleich viel Wert und gleich richtig ist (Perry, 1981, S. 79 f.).

**Position 3: *Multiplicity legitimate but subordinate***

In dieser Position wird die Existenz von Unsicherheit erstmals anerkannt. Es wird nicht mehr zwischen den beiden Kategorien *richtig* und *falsch* sondern zwischen *richtig*, *falsch* und *noch unbekannt* unterschieden. Personen dieser Position sind der Überzeugung, dass für Ansichten, denen aktuell der Status *unbekannt* zuzuordnen ist, einfach noch nicht gesagt werden kann, ob sie richtig oder falsch sind. Dieser Status muss erst noch bestimmt werden. Mit der Unsicherheit wird in dieser Position umgegangen, indem davon ausgegangen wird, dass es eindeutig richtige Wege und Methoden gibt, die *richtige* Lösung/Antwort zu finden und natürlich dementsprechend auch *falsche* Wege oder Methoden (Moore, 2002, S. 20). Die Personen in diesen Positionen richten ihr Lernen dementsprechend danach aus, diese *richtigen* Wege und Methoden zu finden.

**Position 4a: *Multiplicity (diversity and uncertainty) coordinate***

Personen in Position 4 nehmen nach wie vor die dualistische richtig/falsch-Sichweise an. Es wird angenommen, dass ausschließlich Autoritäten den Wahrheitsstatus bestimmter Sachverhalte kennen. Für alle Sachverhalte, zu denen die Autoritäten die richtige Antwort nicht kennen oder aber das Indiviuum die Meinung der Autoritäten nicht kennt, gilt absolute Unsicherheit im Sinne von *wir werden es nie sicher wissen*. Es wird also als unmöglich angesehen festzustellen, ob etwas wahr ist oder bestimmte Ansichten zu gewichten. Meinungen von verschiedenen Personen, die nicht Autoritäten sind, sind demnach völlig arbiträr. Die eigene Meinung wird als die wichtigste empfunden. Aus dieser Sichtweise ergeben sich Einstellungen wie *„Every opinion is right."* und *„Where authorities do not know the answer, any opinion is as good as any other."* (Perry, 1981, S. 84, 86). Hier herrscht die Ansicht, dass eine vollständige Freiheit der Meinungen herrscht oder herrschen sollte. Da jeder das Recht auf eine eigene Meinung hat, die aufgrund der Unsicherheit völlig frei gebildet wird, kann auch keiner die Meinungen anderer Personen als falsch kritisieren (*„They have no right to call me wrong."*, siehe Moore, 2002, S. 20; Perry, 1981, S. 84, 86).

**Position 4b: *Relativism subordinate*** Die Personen, die Position 4b statt oder im Anschluss an 4a durchlaufen, vertrauen in Bereichen mit Unsicherheit stärker auf Autoritäten. Sie bilden sich ihre Meinung also nicht in so vielen Bereichen völlig arbiträr wie Personen in Position 4a (vgl. Perry, 1981, S. 86 f.).

**Kontextabhängiger Relativismus (*Contextual relativism*)**

Die Transition von Position 4 zu 5 stellt die wichtigste Veränderung der individuellen Perspektive innerhalb dieses Modells dar. Die generell dualistische Sichweise auf die Welt, die auch im Multiplizismus noch vorherrscht, auch wenn hier im Vergleich zu den ersten beiden Positionen zunehmend Ausnahmen (im Sinne der Unsicherheit) für bestimmte Situationen anerkannt werden, wandelt sich nun vollständig (Moore, 2002, S. 21). Wissen wird in dieser Position als qualitativ und kontextabhängig angesehen. Diversität in Meinungen, Werten und Einschätzungen, die aus stimmigen Quellen stammen, werden anerkannt. Sie lassen sich mittels Belegen, Logik und Systemen analysieren und vergleichen. So können manche Meinungen als bedeutungslos und unstimmig eingeschätzt werden, während es aber auch Sachverhalte gibt, *„about which reasonable people will resonably disagree“* (Perry, 1981, S. 80).

**Position 5: *Relativism***

Eine relativistische, kontextabhängige Sichtweise auf die Welt zeichnet diese Positionen aus. Richtig-/falsch-Einschätzungen sind nun nur noch die Ausnahme. Personen in dieser Position sind sich ihrer selbst als Konstrukteure von Bedeutung bewusst. Ab dieser Position wird das eigene Einschätzen, Bewerten und Schlussfolgern sowie die Bedeutungszuschreibung als eine zentrale Aufgabe des Lebens erkannt, sowohl in akademischen als auch persönlichen Kontexten (Moore, 2002, S. 21). Das Individuum strebt hauptsächlich danach, die eigenen Fähigkeiten auszuweiten und verschiedene Perspektiven in den unterschiedlichsten Bereichen des Lebens zu erkunden (Perry, 1981, S. 89).

**Selbstverpflichtung/ Verbindlichkeit im Relativismus (*Commitment within relativism*)**

In dieser Entwicklungsphase treffen Personen aktive Entscheidungen zwischen verschiedenen Möglichkeiten (Werte, Meinungen, Politik und persönliche Beziehungen), in vollem Bewusstsein der Relativität des Wissens. Sie verpflichten sich über die Positionen hinweg in steigendem Ausmaß zu Sichweisen und Werten. Die Fähigkeit zu handeln und zu entscheiden wird nun als internal wahrgenommen (Perry, 1981, S. 80). Denken wird als Handeln anerkannt und *etwas zu wissen* stellt immer eine Positionierung des Selbst dar, da somit eine von vielen möglichen Positionen in Relation zu anderen Personen oder Ideen eingenommen wird. Jeder individuelle Akt des Wissens (Denken, Sprechen, Lesen und Schreiben) bedarf, eine bestimmte Perspektive einzunehmen, sich dieser zu verpflichten und sie zu verantworten (vgl. Johnson, 1981 S. 3, siehe Moore, 2002, S. 21).

**Position 6: *Commitment foreseen***

Wissen wird als relativ und kontextabhängig erkannt. Personen dieser Position sind sich darüber bewusst, dass es sich dabei um eine von vielen Möglichkeiten handelt. Ab der Position 6 beginnen Personen in steigendem Ausmaß zu realisieren, dass sie ihre Perspektiven aktiv wählen und sich dieser verpflichten müssen (*commitment*) (Hofer & Pintrich, 1997, S. 91).

**Position 7 bis 9:** ***Evolving commitment***

Die Personen ab Position 7 beginnen zu realisieren, dass sie sich auf eigene Verantwortung in der Unsicherheit des Relativismus der verschiedenen Sichtweisen für eine entscheiden, bzw. eine aktiv wählen müssen. Position 7 geht mit der ersten Selbstverpflichtung einher, beispielsweise zu bestimmten Werten, in Bezug auf soziale Beziehungen oder in Bezug auf eine bestimmte Karriere (Perry, 1981, S. 94). Verantwortung, aktive Auseinandersetzung und Selbstverpflichtung im bestehenden Relativismus kennzeichnen die Personen in diesen Positionen. In steigendem Ausmaß bekennen sie sich zu ihren Werten, Beziehungen und ihrer Identität (Hofer & Pintrich, 1997, S. 91). Wenn die Person weitere Selbstverpflichtungen eingeht (Position 8 und 9), wird es notwendig, Prioritäten bezüglich Energie, Handlung und Zeit zu setzen, um diese auszugleichen (Perry, 1981, S. 95). Die Weiterentwicklung ab Stufe 7 erfolgt bezüglich der Qualität der vorliegenden Selbstverpflichtung. Eine strukturelle Veränderung, wie sie in den Übergängen zwischen den vorherigen Postition vorliegt, findet hier nicht mehr statt (Hofer & Pintrich, 1997, S. 91). Aus diesem Grund sind diese Positionen auch nicht mehr klar zu trennen.

Der beschriebene Entwicklungsverlauf wird überblicksartig in Abbildung 2.1 dargestellt. Die Entwicklung über die ersten fünf Positionen wird von Perry als schrittweise Veränderung des richtig-/falsch-Dualismus charakterisiert. Diese Veränderung erfolgt in dem Versuch, die Diversität in den Meinungen, Erfahrungen und Wahrheiten der Menschen zu berücksichtigen (Perry, 1974, S. 3). Das Erreichen der Position 5 stellt für Individuen einen zentralen Entwicklungsschritt dar (Abb. 2.1). Erstmals wird die relative und kontextabhängige Natur von Wissen erkannt und die eigene Rolle bei der Bedeutungszuschreibung gesehen (Moore, 2002).

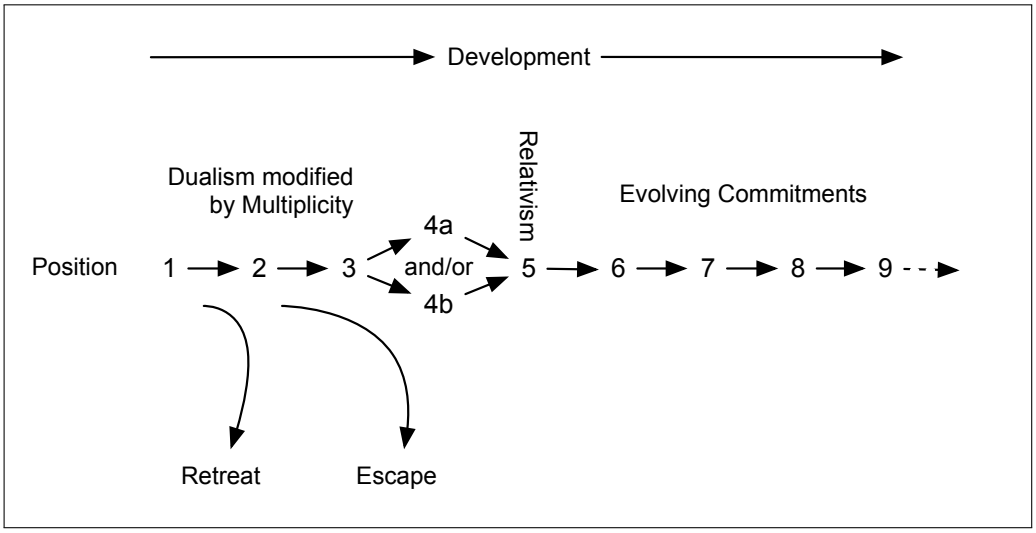


ABBILDUNG 2.1: Epistemologische Entwicklung nach dem *Perry Scheme* (siehe Perry, 1981, S. 80), in eigener Darstellung

Im Bewusstsein der Relativität und Kontextabhängigkeit von Wissen beginnen Individuen ab Stufe 6 aktiv und verbindlich Perspektiven zu wählen und diese zu vertreten. Sie positionieren sich in ihren Ansichten und Werten (Abb. 2.1, S. 29). In der ursprünglichen Konzeptualisierung beschreiben Position 1 bis 5 nach Moore (2002) hauptsächlich den intellektuellen Anteil des Modells im Sinne der systematischen, strukturellen Veränderung in Richtung erweiterter Differenzierung und Komplexität. In Position 6 bis 9 stehen dagegen ethische Fragen, wie Fragen der Identität und Verbindlichkeit in der relativistischen Welt, im Fokus der Entwicklung (ebd.). Perry (1981, S. 80, 90 ff.) identifizierte im Rahmen der Interviewstudien drei Prozesse, die den intellektuellen und ethischen Entwicklungsprozess im Verlauf unterbrechen bzw. diesen am Fortschreiten hindern. Das sogenannte *temporizing* (Abb. 2.1, S. 29) stellt eine Unterbrechung bzw. ein Aufschieben der Entwicklung für ein Jahr oder mehr dar. Die betreffenden Personen warten häufig auf äußere, in der Zukunft liegende Ereignisse, die ihnen eine Entscheidung abnehmen sollen (z.B. *„What do the tests say my interests are?“*, Perry, 1981, S. 90). *Escape* (siebe Abb. 2.1, S.29) bezeichnet eine Entfremdung und das Vernachlässigen von Verantwortungen. Die Personen nutzen die von ihnen nun erkannte Diversität und Relativität des Wissens (Meinungen, Werte etc.) (Position 5) aus, um eine eigene Positionierung in den Meinungen und Werten zu meiden. Unter *retreat* wird der Rückfall von Position 3 in den Dualismus (Position 2) bezeichnet. Die betreffenden Personen meiden Komplexität und Ambivalenz im Wissen, wenden sich für Antworten und Entscheidungen wieder vollständig den Autoritäten zu und lehnen andere Meinungen, Perspektiven und Werte vollständig ab (*„hatred of otherness“*, Perry, 1981, S. 80).

Das *Perry Scheme* betrifft zwei miteinander verbundene Dynamiken: Erstens die Konfrontation und den Umgang mit Diversität und Unsicherheit bezüglich des Lernens und zweitens die Entwicklungsgeschichte der Bedeutungskonstruktion über das Selbst und das Lernen (Moore, 2002, S. 22 f.). Lernende begegnen im Laufe ihrer epistemologischen Entwicklung der Diversität auf zunehmend komplexerem Niveau: Den Anfang bilden multiple Meinungen über Themen und Sachverhalte in den Positionen 1 bis 3. Multiple Kontexte und Perspektiven, aus denen heraus Sachverhalte oder Argumente verstanden und analysiert werden, betreffen die Positionen 4 bis 6. In den Positionen 7 bis 9 werden Lernende mit multipler Selbstverpflichtung/Verbindlichkeit (*commitment*) konfrontiert, über die sie ihre Werte und Identität definieren (ebd.). Während Lernende sich schrittweise mit diesen Aspekten auseinandersetzen, verändert sich der Prozess ihrer Bedeutungskonstruktion und entwickelt sich weiter. Wissen wird als zunehmend unsicher und offen für Interpretation, derer es auch bedarf, angesehen. Das löst einerseits eine Veränderung darin aus, wie die Lernenden die Rolle der Lehrenden sehen. Die Autorität als Quelle der Wahrheit wird über die Zeit zur Ressource mit spezifischer Expertise, die von dieser geteilt wird. Andererseits ändert sich auch die Sichtweise auf die eigene Rolle, als zunächst passive Rezipienten von Fakten bis hin zu aktiv Handelnden, die fähig sind, eigene Argumentationsstrukturen aufzubauen und Wissen zu konstruieren (Moore, 2002, S. 22). Die angenommenen Verbindungen zwischen der epistemologischen Entwicklung zum Lernen von Studierenden wurden von Perry nicht weiter untersucht. Perry (1981) vermutete allerdings bereits Verbindungen zu kognitiven Lernstilen und Lernstrategien.

Perrys Modell, das auf den Interviews mit ausschließlich amerikanischen, männlichen College-Studierenden basiert, bezieht sich dementsprechend auf epistemologische Überzeugungen und Lernen im akademischen Umfeld. Die Erhebung der epistemologischen Überzeugungen der Befragten erfolgte fachdisziplinübergreifend. Es wurde hier also nicht angenommen, dass dieser Prozess für verschiedene Fächer unterschiedlich verläuft oder Personen sich gleichzeitig in verschiedenen Fachdisziplinen in einer unterschiedlichen Phase der Entwicklung befinden könnten.

### 2.2.2 Reflective Judgment Model von King und Kitchener

Auf der Basis von Deweys (1933) Arbeiten zu *reflective thinking* und Perry (1970) entwickelten King und Kitchener das *Reflective Judgment Model* (RJM) (King & Kitchener, 1994). Dewey (1933, 1938), auf den das Konstrukt des reflektierenden Denkens zurückgeht, geht davon aus, dass es sich beim reflektierten Schlussfolgern (*reflective judgments*) um das vorsichtige Abwägen der eigenen Überzeugungen unter Berücksichtigung von stützenden Hinweisen und Belegen handelt. Dieser Prozess wird initiiert, wenn eine Person realisiert, dass im Zusammenhang mit einem Problem, das nicht über formale Logik gelöst werden kann, Kontroversen oder Zweifel vorliegen (King & Kitchener, 2004).

King und Kitchener (1994) übernehmen als Definition epistemologischer Überzeugungen die Konzeptualisierung der *epistemic cognition* von Kitchener (1983), die epistemische Kognition als höchste Ebene der kognitiven Verarbeitung über Kognition und Metakognition einordnet (vgl. Abschnitt 2.3 S. 82). Epistemische Kognition dient nach diesem Verständnis – in Verbindung mit Kognition und Metakognition – dazu, die epistemische Beschaffenheit des Problemlösens zu überwachen. Dies umfasst das Wahrnehmen von eigenen Grenzen des Wissens, der Sicherheit des Wissens und der Kriterien, die im Prozess des Wissenserwerbs eine Rolle spielen. Das RJM befasst sich mit den individuellen epistemologischen Annahmen, die das Verstehen und Schlussfolgern beeinflussen (King & Kitchener, 1994; King et al., 1983; Kitchener & King, 1981). Im Zentrum des Interesses stehen dabei drei Aspekte, die das individuelle Schlussfolgern bezüglich offener, schlecht strukturierter Problemstellungen (z.B. Fragen wie „Are chemical additives harmful?“) betreffen: Die zugrundeliegenden individuellen Annahmen zu Wissen (*epistemic assumptions*), die Art und Weise, wie Individuen schlussfolgern und wie sie ihre eigenen Entscheidungen begründen sowie die Entwicklungssequenz in den Antwortmustern und Einschätzungen zu den Problemstellungen (King & Kitchener, 2004).

Das RJM (King & Kitchener, 1994) entstand auf Basis verschiedener querschnittlicher und längsschnittlicher Interviewstudien mit insgesamt 1700 Personen in der späten Adoleszenz bis ins spätere Erwachsenenalter. Die Erhebung des Entwicklungsstandes reflektierten Schlussfolgerns basierte auf vier komplexen, schlecht strukturierten Problemen (z.B. „Are chemical additives harmful?“) mit je sechs Folgefragen (z.B. "Can scientists ever know for certain?"). Die Befragten wurden gebeten, ihre Sichtweisen zu den Problemen darzustellen und zu begründen. Die Folgefragen bezogen sich explizit auf die individuellen epistemologischen Ansichten der Befragten (King & Kitchener, 1994). Die Interviewtranskripte wurden von trainierten und zertifizierten Ratern eingeschätzt. Die Rater bewerteten die Antworten

der Befragten in Bezug auf die *Beschaffenheit des Wissens* und die *Beschaffenheit der Begründung* (ebd.). Diese beiden übergeordneten Bereiche beinhalteten je 3 Aspekte, die im Rahmen des Ratings eingeschätzt wurden (ebd.). Im ersten Bereich bezog sich das Rating auf die Aspekte: Sichtweise auf Wissen, „richtiges" vs. „falsches" Wissen und Legitimität von Unterschieden in Ansichten. Der zweite Bereich umfasste das Konzept der Begründung, das Nutzen von Belegen und die Rolle von Autorität im Schlussfolgerungsprozess (King & Kitchener, 1994). Das *Reflective Judgment Model* besteht aus sieben qualitativ unterschiedlichen Stufen, die die individuelle Sichtweise und Schlussfolgerungsprozesse bezüglich der schlecht strukturierten Probeme beschreiben. Die Stufen bilden die Entwicklung des reflektierten Denkens ab, d.h. sie führen nach und nach zur Fähigkeit, reflektiert zu schlussfolgern. Die sieben Stufen werden wiederum den folgenden drei übergeordneten Ebenen zugeordnet: *Pre-reflective* (Stufen 1 bis 3), *quasi-reflective* (Stufen 4 und 5) und *reflective* (Stufen 6 und 7) (King & Kitchener, 1994). Jeder der sieben Schritte stellt eine qualitativ unterschiedliche Perspektive epistemologischer Überzeugungen dar (King & Kitchener, 2004, S. 6). Mit jeder Stufe geht eine andere individuelle *Sichtweise auf Wissen* und das *Konzept der Begründung* von Annahmen einher (ebd.). Die Stufen stellen die epistemologischen Überzeugungen von Individuen dar sowie deren Art und Weise, über komplexe (schlecht strukturierte) Sachverhalte zu schlussfolgern (ebd.). Aus der folgenden Beschreibung geht hervor, inwiefern die Annahmen innerhalb jeder einzelnen Stufe untereinander verbunden sind und eine kohärente Einheit bilden (vgl. King & Kitchener, 2004, S. 14 ff.). Zu jeder Stufe wird eine für die entsprechende Entwicklungsstufe indikative Beispielaussage genannt.

**Prä-reflektiert (*Pre-reflective*)**

**Stufe 1**

*Sichtweise auf Wissen:* Es wird angenommen, dass Wissen absolut und konkret ist und keine Abstraktion darstellt. Es kann durch direkte Beobachtung sicher erlangt werden.

*Konzept der Begründung:* Ansichten bedürfen keiner Begründung, da von einer absoluten Übereinstimmung zwischen dem, was als wahr angenommen wird und dem, was wahr ist ausgegangen wird. Alternative Ansichten werden nicht wahrgenommen.

*„Ich weiß, was ich gesehen habe."*

**Stufe 2**

*Sichtweise auf Wissen:* Es wird angenommen, dass Wissen absolut sicher oder aber sicher, jedoch aktuell nicht verfügbar ist. Wissen kann direkt über die Sinne (mittels direkter Beobachtung) oder über Autoritätspersonen erlangt werden.

*Konzept der Begründung:* Annahmen sind ungeprüft und unbegründet oder aber werden über ihre Übereinstimmung mit der Annahme einer Autoritätsperson (wie einer Lehrkräft oder Elternteil) begründet. Für die meisten Themen/Fragen wird eine richtige Antwort (Lösung) angenommen, sodass dem Individuum kaum oder gar kein Konflikt in der Entscheidungsfindung über umstrittene Fragen entsteht.

*„Wenn es in den Nachrichten kommt, muss es wahr sein."*

**Stufe 3**

*Sichtweise auf Wissen:* Es wird angenommen, dass Wissen absolut sicher oder aber temporär unsicher ist. In Bereichen mit zeitweiliger Unsicherheit können nur persönliche Annahmen gewusst werden, bis das absolute Wissen erlangt wird. In Bereichen mit absoluter Sicherheit wird Wissen von Autoritäten erlangt.

*Konzept der Begründung:* In Bereichen, in denen sichere Antworten existieren, werden Annahmen gerechtfertigt, indem auf die Sichtweisen von Autoritäten referenziert wird. In Bereichen, in denen es keine sicheren Antworten gibt, werden Annahmen gerechtfertigt, indem auf sie als persönliche Meinung referenziert werden, da in diesen Bereichen die Verbindung zwischen Evidenz und Annahme unklar ist.

*„Wenn es Belege gibt, die alle Menschen von der einen oder anderen Seite überzeugt, dann handelt es sich dabei um Wissen; bis dahin handelt es sich nur um eine Vermutung."*

**Quasi-reflektiert (*Quasi-reflective*)**

**Stufe 4**

*Sichtweise auf Wissen:* Wissen ist unsicher und Wissensansprüche sind dem jeweiligen Individuum zu eigen, da situationale Faktoren (wie z.B. fehlerhafte Berichte von Informationen, über Zeit verlorene Informationen oder ungleicher Zugang zu Informationen) dafür sorgen, dass Wissen immer mit Ambiguität behaftet ist.

*Konzept der Begründung:* Annahmen werden über Gründe und Belege gerechtfertigt, allerdings sind die Argumente und die Wahl der Belege idiosynkratisch (z.B. Wahl eines Beleges, der zur eigenen Annahme passt).

*„Ich wäre eher geneigt an die Evolution zu glauben, wenn es Beweise dafür gäbe. Das ist genau wie mit den Pyramiden: Ich glaube nicht, dass wir es je wissen werden. Wen sollte man fragen? Keiner war dort."*

**Stufe 5**

*Sichtweise auf Wissen:* Wissen ist kontextabhängig und subjektiv, weil es durch die Wahrnehmung einer Person und deren Kriterien der Einschätzung gefiltert wird. Nur Interpretationen von Belegen, Geschehnissen und Themen sind erfahrbar.

*Konzept der Begründung:* Annahmen werden innerhalb eines gewissen Kontexts über die Regeln der Überprüfung, die für diesen spezifischen Kontext gelten und über kontextspezifische Interpretationen der Evidenz begründet. Bei bestimmten Annahmen wird davon ausgegangen, dass sie kontextspezifisch sind oder gegen andere Interpretationen abgewogen werden müssen. Dies kompliziert (und verzögert manchmal) Schlussfolgerungen.

*„Personen denken unterschiedlich und deshalb gehen sie ein Problem unterschiedlich an. Andere Theorien könnten genauso zutreffen, wie meine eigene, nur auf anderen Belegen basieren."*

**Reflektiert (*Reflective*)**

**Stufe 6**

*Sichtweise auf Wissen:* Wissen wird in individuellen Schlussfolgerungen über schlecht strukturierte Probleme konstruiert. Dies geschieht auf der Basis von Informationen aus verschiedensten Quellen. Interpretationen, die auf der Auswertung von kontextübergreifenden Evidenzen und den überprüften bzw. abgewogenen Meinungen von verlässlichen Personen basieren, können gewusst werden.

*Konzept der Begründung:* Annahmen werden einerseits begründet, indem ein Vergleich von Evidenzen und Meinungen aus verschiedenen Perspektiven zum Thema über verschiedene Kontexte vorgenommen wird. Andererseits werden sie begründet, indem Lösungen konstruiert werden, die mittels Kriterien überprüft werden, wie z.B. dem Gewicht eines Belegs, dem Nutzen der Lösung oder dem pragmatischen Handlungsbedarf.

*„Es ist in diesem Leben sehr schwer, sich bei etwas sicher zu sein. Es gibt Abstufungen der Sicherheit. Man kann an einen Punkt kommen, an dem man sich sicher genug ist, um sich eine persönliche Haltung zu einem Thema zu bilden."*

**Stufe 7**

*Sichtweise auf Wissen:* Wissen ist das Ergebnis eines Prozesses der angemessenen Prüfung, in dem Lösungen zu schlecht strukturierten Problemen konstruiert werden. Die Angemessenheit dieser Lösungen wird danach geprüft, welche im Hinblick auf die betreffenden Evidenzen am logischsten oder wahrscheinlichsten ist. Dies muss erneut überprüft werden, wenn neue Belege, Perspektiven oder Mittel der Prüfung zur Verfügung stehen.

*Konzept der Begründung:* Annahmen werden probabilistisch bewertet. Dies geschieht auf Basis der verschiedensten interpretativen Aspekte, wie das Gewicht der Belege, der Erklärungswert einer Interpretation und das Risiko, das fehlerhafte Schlussfolgerungen mit sich bringen, die Konsequenzen alternativer Schlussfolgerungen und die Beziehung zwischen all diesen Faktoren. Schlussfolgerungen werden darüber verteidigt, dass sie das vollständigste, plausibelste oder überzeugenste Verständnis eines Themas/Frage auf Basis der verfügbaren Belege repräsentieren.

*„Ein Argument kann danach bewertet werden, wie gut durchdacht die entsprechende Position ist, welche Arten von Schussfolgerungen und Belegen genutzt werden, um es zu stützen und wie stimmig die Argumentationsweise einer Person in diesem Bereich im Vergleich zu anderen Bereichen ist."*

Reflektiertes Schlussfolgern wird von King und Kitchener (1994) im Bereich der intellektuellen Entwickung angesiedelt. Das Konstrukt des *kritischen Denkens* betrachten sie als verwandt zum *reflektierten Schlussfolgern*, grenzen es aber insofern ab, als dass sich ersteres eher auf das Lösen von geschlossenen, gut strukturierten Problemen beziehe. Zwischen dem reflektierten Schlussfolgern und der psychologischen Entwicklung fanden King und

Kitchener (1994) moderate Zusammenhänge. In einem Review konnten King und Kitchener (1994) zeigen, dass die Entwicklung des reflektierten Schlussfolgerns mit der Entwicklung anderer Aspekte der kognitiven Entwicklung (*verbal aptitude, formal operation, academic ability, critical thinking*) sowie der moralischen Entwicklung und Identitätsentwicklung zwar zusammenhängt, aber sich doch davon trennen lässt. Darüber hinaus sagt die Entwicklung des reflektierten Schlussfolgerns sehr stark die Toleranz gegenüber Diversität vorher (Guthrie, King & Palmer, 2000).

Das RJM liefert insofern einen bedeutenden Beitrag zum Verständnis der epistemologischen Überzeugungen, als dass es die von Perry hypothetisch angenommenen und eher sehr ungenau beschriebenen oberen Stufen ausdifferenziert (Hofer & Pintrich, 1997; Moore, 2002). Darüber hinaus erlaubt es der Erhebungsansatz der epistemologischen Entwicklung über die komplexen Probleme den Autoren, einen Bereich der intellektuellen Entwicklung zu erfassen, der über die Forschung zum kritischen Denken nicht abgedeckt wird. Allerdings wird anhand der vielen anderen Konzeptionen auch deutlich, dass epistemologische Überzeugungen sich nicht nur über das komplexe Problemlösen erfassen lassen. Hofer und Pintrich (1997) merken hierzu an, dass Personen schließlich auch im alltäglichen Leben Annahmen zu Wissen und dem Wissensprozess haben, die über alltägliche Bildungskontexte aktiviert werden und die das Lernen grundlegend beeinflussen. Die Antworten der Befragten zu hypothetischen Problemen mögen wenig darüber aussagen, wie die Überzeugungen in alltäglichen Lernerfahrungen beschaffen sind (Hofer & Pintrich, 1997). Auch findet nur ein geringer Anteil der Entwicklung überhaupt in dem untersuchten Zeitraum (College) statt. Es bleibe unklar, wie genau Bildung die epistemologische Entwicklung beeinflusse und wie sich die epistemologische Entwicklung in Kontexten entwickeln (ebd.). Wichtig ist festzuhalten, dass King und Kitchener mit ihrem Modell explizit keinen Anspruch auf kulturübergreifende Gültigkeit erheben (King & Kitchener, 2004).

Auch wenn das Modell über die bei der Erhebung eingesetzten Problemstellungen verschiedene Kontexte erzeugt, werden diese bezüglich der epistemologischen Überzeugungen nicht berücksichtigt. Nach Zinn (2013) wird in dieser Konzeption von einer Abhängigkeit der epistemologischen Überzeugungen von der betreffenden Disziplin oder dem Thema nicht ausgegangen, sie wird nicht berücksichtigt.

### 2.2.3 Development of Epistemological Understanding von Kuhn

Kuhn und Weinstock (2002, S. 134) beschreiben die Entwicklung epistemologischer Überzeugungen in ihrem Entwicklungsmodell (Kuhn, 1991) als *„theory-in-action“*, im Sinne von impliziten Theorien zu Wissen und zum Wissensprozess, die die alltägliche Wissenskonstruktion und -begründung beeinflussen. Sie zielen darauf ab, das epistemologische Verstehen vom Kindergartenalter bis ins Erwachsenenalter zu erklären und beziehen sich dabei nicht nur auf den akademischen Bereich sondern auch den Lebensalltag. Kuhn (1991) legte ihren Fokus auf Denkprozesse bezüglich schlecht strukturierter Probleme im Alltag. Wie

schon bei King et al. (1983) und King und Kitchener (1994) erfolgt die Erhebung der epistemologischen Überzeugungen über die individuellen Schlussfolgerungsprozesse zu spezifischen Problemen des Alltags, statt offene Interviews oder standardisierte Fragebögen einzusetzen (Kuhn & Weinstock, 2002). Ihr Forschungsdesign umfasste je vier bezüglich Geschlecht und Bildungsniveau (College, kein College) balancierte Altersgruppen (Teenager, 20er, 40er, 60er) mit je ca. 40 Probanden (Kuhn, 1991).

Das Modell Kuhns gehört zu den Entwicklungsstufenmodellen und unterscheidet zwischen vier epistemologischen Sichtweisen (*epistemological views*, Kuhn, 1991; Kuhn & Pearsall, 2000; Kuhn & Weinstock, 2002): Realist, Absolutist, Multiplizist und Evaluatist. In Tabelle 2.1 sind die vier Stufen überblicksartig dargestellt. Der generelle Ablauf stimmt mit den anderen Entwicklungsstufenmodellen überein (Kuhn & Weinstock, 2002), auf die im Rahmen der Forschungsentwicklung (erste Forschungsphase) verwiesen wurde. Die Stufen weisen so große Parallelen zu Perrys sowie King und Kitcheners Modellen auf, dass Hofer und Pintrich (1997) sie später als verdichtete Versionen derselben beschreiben. Personen auf diesen vier Stufen des epistemologischen Verstehens von Kuhn unterscheiden sich jeweils qualitativ in Bezug auf ihr zugrundeliegendes Verständnis von *Aussagen*, *Realität*, *Wissen* und *kritischem Denken* (siehe Tab. 2.1).

TABELLE 2.1: Stufen epistemologischen Verstehens nach Kuhn und Pearsall (2000, S. 311)

| Stufe | Aussagen | Realität | Wissen | Kritisches Denken |
|---|---|---|---|---|
| Realist | Aussagen sind KOPIEN der externen Realität | Realität ist direkt erfahrbar. | Wissen hat seinen Ursprung in externer Quelle und ist sicher. | Kritisches Denken ist nicht nötig. |
| Absolutist | Aussagen sind FAKTEN, die korrekt oder inkorrekt in ihrer Repräsentation der Realität sind (Möglichkeit falscher Überzeugungen) | Realität ist direkt erfahrbar. | Wissen hat seinen Ursprung in externer Quelle und ist sicher. | Kritisches Denken dient dazu, Behauptungen mit der Realität abzugleichen, um ihre Richtigkeit oder Falschheit einzuschätzen. |
| Multiplizist | Aussagen sind MEINUNGEN, die durch die Person, die über sie verfügt, frei gewählt werden und sich nur dieser gegenüber zu verantworten haben. | Realität ist nicht direkt erfahrbar. | Wissen wird durch den menschlichen Verstand entwickelt und ist unsicher. | Kritisches Denken ist irrelevant. |
| Evaluatist | Behauptungen sind EINSCHÄTZUNGEN, die evaluiert und mittels Kriterien des Argumentierens und der Evidenz verglichen werden können | Realität ist nicht direkt erfahrbar. | Wissen wird durch den menschlichen Verstand entwickelt und ist unsicher. | Kritisches Denken dient der zuverlässigen Einschätzung und fördert das Verstehen. |

Die epistemologische Entwicklung des Verstehens umfasst die Koordination zwischen subjektiven und objektiven Sichtweisen auf das Wissen (Kuhn & Park, 2005). Bei den vier Stufen handelt es sich um eine Sequenz des qualitativ unterschiedlichen Verständisses dessen, was es bedeutet ein Wissensurteil zu fällen. Dieses Verständnis reicht von Wissensurteilen als Kopien der Realität, über Fakten, Meinungen bis hin zu Urteilen (Kuhn & Weinstock, 2002, S. 125, siehe Tab. 2.1, S. 36). Nach dieser Sichtweise prägen die epistemologischen Überzeugungen nicht nur, wie Individuen eine Aufgabe angehen werden, sondern sie werden gleichzeitig auch selbst durch den Kontext (z.B. Aufgabenbedingungen, Inhaltsbereich, soziale und kulturelle Werte und Praktiken) aktiviert und geformt (Kuhn et al., 2008; Kuhn & Pearsall, 2000; Kuhn & Weinstock, 2002; Kuhn et al., 2013; Tabak & Weinstock, 2008). Kuhn (1991) untersuchte die Verbindungen zwischen der Stufe der epistemologischen Entwicklung und der domänenspezifischen Argumentationsfähigkeit. Sie identifizierte die folgenden drei Fertigkeiten der Argumentation: Entwicklung von echter Evidenz, alternativen Theorien und Gegenargumenten. Personen mit einer evaluativen Perspektive auf Wissen setzten tatsächlich mehr dieser drei Argumentationsfertigkeiten ein, als Personen auf niedrigerer Stufe der epistemologischen Entwicklung (ebd.).

Kuhns (1991) besonderer Beitrag zur Forschung im Themenbereich besteht darin, die Verbindung zwischen epistemologischen Überzeugungen und dem Schlussfolgern (reasoning) herzustellen. Auch über die breitere Stichprobenbildung bezüglich Geschlecht, Bildungsniveau und Altersgruppen der Probanden und damit einhergehend der Verortung epistemologischer Überzeugungen im nicht-akademischen Kontext hebt sich die Forschung Kuhns (1991) von den übrigen Forschungsbeiträgen zur epistemologischen Entwicklung ab (vgl. Hofer & Pintrich, 1997). Es werden inhaltliche Domänenunterschiede in den epistemologischen Überzeugungen angenommen, wobei der Begriff Domäne sich bei Kuhn (1991) auf den Urteilsbereich bezieht, im Sinne von Urteilen über Wahrheit, Urteilen über persönlichen Geschmack, Ästhetik oder Werte. Disziplinspezifische Unterschiede in epistemologischen Überzeugungen werden nach Zinn (2013, S. 62) in dieser Konzeption demnach nicht berücksichtigt.

### 2.2.4 Epistemological Belief System von Schommer

Schommers (später Schommer-Aikins) Forschungsprogramm (Schommer, 1990, 1993a; Schommer et al., 1992) zielte einerseits darauf ab, mehr über Beschaffenheit der epistemologischen Überzeugungen herauszufinden und andererseits ihre Einflüsse auf das Verstehen und Lernen in akademischen Disziplinen aufzudecken. Mit ihrem Modell werden epistemologische Überzeugungen als System von zunächst fünf mehr oder weniger unabhängigen Überzeugungsdimensionen angesehen. Die Beschreibung der Dimensionen als *mehr oder weniger unabhängig* stammt von Schommer selbst. Sie meint damit, „that an individual is not necessarily at the same level of sophistication across all the epistemological dimensions“ (Schommer-Aikins, 2008, S. 304). Die Entwicklung in den verschiedenen Dimensionen kann also asynchron verlaufen (Schommer, 1994; Schommer-Aikins, Bird & Bakken, 2010). Schüler können beispielsweise der Überzeugung sein, dass Wissen größtenteils komplex ist

und gleichzeitig auch, dass Wissen überwiegend sicher ist (Schommer-Aikins et al., 2010, S. 32). Dieser Ansatz der epistemologischen Überzeugungen als *Belief System* unterscheidet sich nach Darstellungen Schommers in sechs zentralen Aspekten von den bis dahin üblichen entwicklungspsychologischen Ansätzen (vgl. Schommer-Aikins, 2004, S. 20 f.):

1. Voneinander getrennte Überzeugungsdimensionen können nun erfasst und auch einzeln untersucht werden.
2. Im Modell wird berücksichtigt, dass sich die verschiedenen Überzeugungsbereiche unabhängig voneinander entwickeln können.
3. Im Gegensatz zu den Entwicklungsstufenmodellen werden hier Überzeugungen zum Lernen in das Modell integriert.
4. Es wird berücksichtigt, dass ein Bedarf nach der Balance zwischen den gegenüberliegenden Ausprägungen von epistemologischen Überzeugungen vorliegt (need of balance).
5. Es wird für die epistemologischen Überzeugungen eine einheitliche Nomenklatur eingeführt.
6. Die quantitative Erhebung epistemologischer Überzeugungen wird eingeführt.

Schommer (1990) nahm in ihrer ersten Konzeptualisierung a priori die fünf Dimensionen *stability of knowledge*, *structure of knowledge* und *source of knowledge* sowie *speed of knowledge acquisition* und *ability of knowledge acquisition* an. Die Idee einer Dimension zur *Fähigkeit des Wissenserwerbs* geht auf Dweck und Leggetts (1988) Arbeit zur Beschaffenheit der Intelligenz zurück. Von Schoenfelds Forschung zu Überzeugungen über Mathematik (Schoenfeld, 1983, 1985) stammt der Ansatz, getrennte Überzeugungsdimensionen und solche zum Lernen anzunehmen (siehe Schommer, 1990; Schommer-Aikins, 2004). Die übrigen drei Dimensionen zur Struktur, Quelle und Stabilität basieren auf der Forschung von Perry (1968) (Schommer, 1990). Der Hinweis M. P. Ryans (1984), dass objektivere Fragebögen eine Möglichkeit zur Erfassung epistemologischer Überzeugungen darstellen könnten, inspirierte Schommer (1990) zu dieser Erhebungsmethode (Schommer-Aikins, 2004, S. 20). Im Rahmen des Modells wird für die epistemologischen Überzeugungen insofern eine einheitliche Nomenklatur eingeführt, als dass Schommer in dieses Forschungsfeld den Begriff der *Beliefs* (*Überzeugungen*) einbringt. Aus dem Bereich der Überzeugungen (Abelson, 1986; Alexander & Dochy, 1994; Pajares, 1992) übernimmt Schommer für die epistemologischen Überzeugungen die Aspekte „affect, limited adherence to logic, difficulty in changing, and a powerful influence on thinking“ (Schommer-Aikins, 2004, S. 21). Epistemologische Überzeugungen sind nach dieser Sichtweise oft unbewusst, relativ stabil, weil schwer veränderbar und haben einen großen Einfluss auf Denkprozesse (ebd.). Hier lassen sich Parallelen zum Traitparadigma der Persönlichkeitspsychologie erkennen (Asendorpf, 1999). Schommers *Epistemological Belief System* wird als persönlichkeitspsychologischer Ansatz betrachtet. Die mehrdimensionale Konzeptualisierung und Erhebung mittels Items sowie die Auswertung mittels faktorenanalytischer Verfahren ermöglicht es, individuelle Unterschiede in den Ausprägungen auf den

verschiedenen Dimensionen zu betrachten. Nach dem Vorbild von Persönlichkeitseigenschaften können die individuellen Ausprägungen in epistemologischen Überzeugungen ins Verhältnis zu einer Referenzpopulation gesetzt werden. Die individuelle Ausprägung wird verglichen mit der eines durchschnittlichen Mitglieds der Referenzpopulation (Klopp, 2014, S. 15).

Im folgenden werden die Pole der fünf Überzeugungsdimensionen kurz beschrieben (nach Schommer-Aikins, 2004, S. 20). Die erst genannte Sichtweise entspricht jeweils der, die Schommer (1990) als die *naive* epistemologische Überzeugung betrachtet, während die andere die *sophistizierte* darstellt. Es handelt sich bei den Dimensionen also jeweils um ein Kontinuum mit zwei Polen. Die gegenüberliegenden Pole entsprechen den beiden extremsten Ausprägungen epistemologischer Überzeugungen.

**stability of knowledge:**
*Wissen ist unveränderbar/absolut* bis hin zu *Wissen ist vorläufig*.

**structure of knowledge/organisation of knowledge:**
*Wissen besteht aus isolierten, kleinen Bestandteilen* bis hin zu *Wissen besteht aus integrierten Konzepten*.

**source of knowledge:**
*Wissen liegt in allwissenden Autoritäten und wird durch diese vermittelt* bis hin zu *Wissen hat seinen Ursprung in Begründungen und empirischer Evidenz*.

**speed of learning:**
*Es wird entweder schnell oder gar nicht gelernt* bis hin zu *Lernen erfolgt sukzessive*.

**ability to learn:**
*Die Fähigkeit zu lernen ist angeboren* bis hin zu *die Fähigkeit zu lernen ist steigerungsfähig*.

Zu diesen fünf Dimensionen entwickelte Schommer (1990) zunächst insgesamt 63 Aussagen zu Lernen und Wissen, die zwölf Item-Subsets angehören. Die Zuordnung der Subsets zu den Dimensionen sowie die jeweilige Itemanzahl ist Tabelle 2.2 auf Seite 40 zu entnehmen. 35 Items entsprachen in ihrer Formulierung einer sogenannten *naiven* Ausprägung epistemologischer Überzeugung und 28 Items der *fortgeschrittenen* (Schommer, 1990, S. 499). Die Studierenden gaben den Grad der Zustimmung zu diesen Aussagen auf einer Likertskala (von 1 *strongly disagree* bis 5 *strongly agree*) an (ebd.). Das Erhebungsinstrument wird *Schommmers Epistemological Questionnaire* genannt. Es wird in der Literatur uneinheitlich teilweise mit SEQ, EQ oder EBQ abgekürzt. Die Items sind bei Schommer (1998) vollständig abgedruckt. In Schommers ursprünglichem Modell werden domänenspezifische Unterschiede in den Überzeugungen zwischen verschiedenen akademischen Disziplinen nicht berücksichtigt. Es handelt sich im Rahmen des SEQ demnach um eine domänen-übergreifende Erhebung (siehe SEQ, Schommer, 1998).

TABELLE 2.2: Zuordnung der 12 Subsets zu den 5 Subskalen des SEQ

| Dimensionen | Subsets | Itemanzahl |
|---|---|---|
| structure/simple knowledge | seek single answers | 11 |
| | avoid integration | 8 |
| certain knowledge | avoid ambiguity | 5 |
| | knowledge is certain | 6 |
| authority/ source of knowledge | depend on authority | 4 |
| | don't criticize authority | 6 |
| controll/innate abbility | success is unrelated to hard work | 4 |
| | ability to learn in innate | 4 |
| | can't learn how to learn | 5 |
| speed/quick learning | learning is quick | 5 |
| | learn first time | 3 |
| | concentrate effort is a waste of time | 2 |

Eine faktorenanalytische Prüfung auf Basis der zwölf Item-Subsets von Schommers Dimensionen (Schommer, 1990) führte zu vier Faktoren, die die folgenden vier Dimensionen abbilden: *Fixed Ability* [2], *Quick Learning*, *Simple Knowledge*, *Certain Knowledge*. Die Benennung der vier Dimensionen entspricht jeweils der *naiven* Ausprägung auf der entsprechenden Dimension. Diese vier Dimensionen konnten in späteren Studien repliziert werden (Jehng et al., 1993; Schommer, 1993a, 1998; Schommer et al., 1992). Die fünfte angenommene Dimension *Quelle des Wissens* konnte faktorenanalytisch (auf Basis der vorhandenen Subsets und nicht Items) von Schommer nicht bestätigt werden. Stattdessen luden die für diese Dimension ursprünglich vorgesehenen Subsets auf anderen Faktoren (Schommer, 1990, 1993a, 1993b; Schommer et al., 1992).

**Certain Knowledge**

Diese Dimension umfasst die beiden Pole mit Überzeugungen „Wissen ist absolut" bis hin zu „Wissen ist vorläufig und entwickelt sich". Sie besteht aus dem Item-Subset *Knowledge Is Certain* (Beispielitem: „The only thing that is certain is uncertainty itself."). Zwar wurden für diesen Faktor noch weitere Item-Subsets entwickelt, diese werden hier jedoch nicht aufgeführt, da sie nicht konsistent auf diesen Faktor laden (Schommer, 1998).

**Simple Knowledge**

Diese Dimension umfasst die beiden Pole mit Überzeugungen „Wissen besteht aus isolierten, klar trennbaren Bestandteilen" bis hin zu „Wissen besteht aus integrierten Konzepten". Ihr werden die drei Subsets *Avoid Ambiguity* (Beispielitem: „I don't like movies that don't have an ending."), *Seek Single Answers* (Beispielitem: „Most words have one clear meaning.") und *Avoid Integration* (Beispielitem: „When I study I look for specific facts.") zugeordnet (Schommer, 1998).

[2] Die Bezeichnung *Fixed Ability* geht auf Dweck und Leggetts (1988) zurück.

**Quick Learning**

Diese Dimension umfasst die beiden Pole mit Überzeugungen „Es wird entweder schnell oder gar nicht gelernt“ bis hin zu „Lernen erfolgt sukzessive“. Sie besteht aus dem Subset *Learning is quick* (Beispielitem: „Successful students learn things quickly.“) (Schommer, 1998).

**Fixed Ability**

Diese Dimension umfasst die beiden Pole mit Überzeugungen „Die Fähigkeit zu lernen ist angeboren“ bis hin zu „die Fähigkeit zu lernen ist steigerungsfähig“, d.h. es kann gelernt werden, wie man lernt. Sie umfasst die Subsets *Cant't learn how to learn* (Beispielitem: „Self-help books are not much help.“), *Success is unrelated to hard work* (Beispielitem: „The really smart students don't have to work hard to do well in school.“), *Learn the first time* (Beispielitem: „Almost all the information you get from a textbook you will get during the first reading.“) und *Ability to learn is innate* (Beispielitem: „An expert is someone who has a special gift in some area.“). Das letzte dieser Subsets lädt allerdings nicht konsistent auf diesem Faktor, sondern teilweise auch auf *Quick Learning* (Schommer, 1998).

Auch direkte Replikationsstudien (z.B. von Bendixen et al., 1994; Clarebout et al., 2001; DeBacker & Crowson, 2008; Klopp, 2014; Paulsen, 1998) sowie Studien, die den SEQ in abgewandelter, gekürzter oder übersetzter Form einsetzten (z.B. Cole & Goetz, 2000; Qian & Alvermann, 1995) fanden von den theoretisch postulierten fünf angenommenen Dimensionen abweichende Faktorstrukturen. Die Studien identifizierten uneinheitlich teilweise zwei, drei oder vier Faktoren. Dies spricht allerdings nicht gegen Schommers theoretisches Modell, sondern vielmehr gegen die Eignung des Instruments.

Schommer schlägt später vor, epistemologische Überzeugungen als Häufigkeitsverteilungen und nicht als Kontinuum zu verstehen (Schommer, 1994; Schommer-Aikins & Hutter, 2002). Ein Individuum kann also annehmen, dass der Großteil des eigenen Wissens veränderbar ist und gleichzeitig der Überzeugung sein, dass es einen kleinen Anteil an unveränderlichem Wissen gibt. Diese Person würde dann dazu neigen, anzunehmen und so zu handeln, als wäre Wissen veränderbar. Gleichzeitig würde sie für einen Teil ihres Wissens Stabilität annehmen (Schommer-Aikins, 2004, S. 21). Schommers Annahme zu *Need of Balance*, d.h. der Notwendigkeit in Bezug auf die eigenen individuellen Überzeugungen eine Balance herzustellen, ist eng mit diesem vorherigen Vorschlag verbunden. Schließlich könnten sich extreme Überzeugungen in beide Richtungen als problematisch herausstellen. Eine extreme Ausprägung in der Überzeugung, dass Wissen in jedem Fall unsicher und veränderlich ist, würde eine Person schließlich genauso einschränken, wie die extreme Überzeugung, dass Wissen absolut sicher ist. Im ersten Fall wäre die Person vielleicht nicht fähig, eine eigene Sichtweise zu entwickeln und zu vertreten, der zweite Fall könnte dazu führen, dass die Person ablehnt, Neues zu lernen. Es herrscht dementsprechend die Notwendigkeit, ein Gleichgewicht herzustellen und damit geht auch die Annahme einher, epistemologische Überzeugungen als Häufigkeitsverteilungen zu betrachten (ebd.). Der Grad der Sophistiziertheit epistemologischer Überzeugungen drückt sich nach dieser neueren Ansicht nicht

mehr ausschließlich in den extremen Überzeugungen aus, die dem einen Pol des Kontinuums entsprechen, sondern vielmehr in einer gewissen Balance zwischen den Polen. Ein Individuum mit vorherrschenden *sophistizierten* Überzeugungen zur Sicherheit (veränderliches Wissen) und Komplexität (integrierte Konzepte) des Wissens sowie Lernfähigkeit (erlernbar) und -geschwindigkeit (sukzessive) mit geringen Anteilen an Überzeugungen aus dem gegenüberliegenden Pol ist demnach anpassungs- und handlungsfähiger (Schommer-Aikins, 2004).

Der zentrale Beitrag Schommers (später Schommer-Aikins) besteht nach Einschätzung Hofer und Pintrichs (1997) darin, ein System von Dimensionen vorzuschlagen und die empirische Prüfung der Existenz von verschiedenen Überzeugungsdimensionen anzuregen. Darüber hinaus leitete sie die Forschung zur Verbindung zwischen den epistemologischen Überzeugungsdimensionen mit dem akademischem Lernen und Leistung ein (ebd.). Sie konnte zeigen, dass die epistemologischen Überzeugungen Verbindungen mit Verstehen (Schommer, 1990), Leistung und Lernstrategienutzung (Schommer et al., 1992) aufweisen. Die Verbindung zwischen epistemologischen Überzeugungen, Lernstrategien und Lernleistung stellt Schommer-Aikins (2004) in einer Weiterentwicklung ihres ursprünglichen Modells (Schommer, 1990), dem *Embedded Systemic Model* dar. Schommer-Aikins (2004) konzeptualisiert hier epistemologische Überzeugungen als „system-among-other-systems“, da sie nicht in einem Vakuum funktionieren, sondern „[...] leaners' thoughts, actions, or motivations represent the convergence of multiple systems.“ (Schommer-Aikins, 2004, S. 23). Das Modell wird in Kapitel 2.8.1 im Zusammenhang mit der Darstellung der Beziehung von epistemologischen Überzeugungen und dem Lehr-Lernprozess vorgestellt (siehe Abbildung 2.9, S. 88).

Allerdings erfuhr Schommers Konzeptualisierung auch deutliche Kritik im Hinblick auf ihre Definition des Konstrukts epistemologischer Überzeugungen. Nach der Argumentation von Hofer und Pintrich (1997) handelt es sich bei *Fixed Ability* und *Quick Learning* um Dimensionen, die nicht zum Konstrukt der epistemologischen Überzeugungen gehören, da sie sich nicht auf das Wissen beziehen. Die Überzeugung eines Lernenden dazu, ob Lernen schnell oder gar nicht bzw. langsam vonstatten geht (*Quick Learning*), ist nach (Hofer & Pintrich, 1997) vielmehr Ausdruck der wahrgenommenen Aufgabenschwierigkeit anstatt eine Überzeugung zur Natur des Wissens. Auch die Dimension *Fixed Ability* lässt sich demnach nicht in den Kernbereich epistemologischer Überzeugungen einordnen. Die individuelle Überzeugung darüber, ob die Fähigkeit etwas zu lernen angeboren ist oder aber gelernt werden kann, betrifft vielmehr „the nature of intelligence as a personal, psychological trait of an individual“ (Hofer & Pintrich, 1997, S. 109). Hofer und Pintrich (1997, S. 109) sehen darin, dass mittels des SEQ verschiedene Konstrukte (epistemologische Überzeugungen, Überzeugungen zu Intelligenz oder Fähigkeiten) erhoben würden, die Ursache dafür, dass sich die Dimension *Fixed Ability* unabhängig zu den anderen Dimensionen verhält, während Schommer diese Unabhängigkeit dazu veranlasse ihre Dimensionen als „more or less independent“ (Schommer, 1990, S. 498) zu beschreiben. Weitere starke Kritikpunkte betreffen die Validität der Erhebung mittels des SEQ. Es ist nach Hofer und Pintrich (1997, S. 109 f.) zu bezeifeln, dass Items wie „Nothing is certain but death and taxes“ oder „Self-help

books are not much help.“ relevant, repräsentativ für den Zielbereich und indikativ für das Zielkonstrukt sind, wie jedoch nach Messick (1989) für Items gefordert wird.

### 2.2.5 Epistemological Theories von Hofer und Pintrich

Wie oben bereits kurz beschrieben, entstand die Konzeptualisierung epistemologischer Überzeugungen als *Persönliche Theorien* von Hofer und Pintrich (1997) auf Basis ihres Reviews über die Konzeptionen epistemologischer Überzeugungen von Perry (1970), Belenky et al. (1986), Baxter Magolda (1992), King und Kitchener (1994), Kuhn (1991) und Schommer (1990). Hofer und Pintrich (1997) identifizierten in den sechs Modellen eine Schnittmenge bezüglich des Verständnisses des Konstrukts epistemologischer Überzeugungen. Die Modelle weisen also teilweise sehr große Gemeinsamkeiten bezüglich der inhaltlichen Aspekte epistemologischer Überzeugungen auf. In Abbildung 2.2 auf S. 44 wird die Gegenüberstellung der Komponenten epistemologischer Überzeugungen aus den von Hofer gereviewten Modellen dargestellt (siehe Hofer & Pintrich, 1997, S. 113 ff). Wie sich hier zeigt, finden sich in den verschiedenen Modellen insgesamt die epistemologischen Überzeugungsdimensionen: *Sicherheit des Wissens*, *Einfachheit des Wissens*, *Quelle des Wissens*, *Rechtfertigung des Wissens*, *Beschaffenheit des Lernens* und *Beschaffenheit der Intelligenz*. Die in den Konzeptionen jeweils angenommenen Ausprägungen der Dimensionen stehen ober- und unterhalb des schwarzen Pfeils (siehe 2.2, S. 44). Alle sechs oben genannten Modelle verfügen beispielsweise jeweils über eine Definition zu Überzeugungen zur Quelle des Wissens. Auch die Beschreibungen der Ausprägungen, die diese Überzeugungen annehmen können, ähneln sich über die Modelle hinweg sehr stark. Auch Überzeugungen zur Sicherheit, Einfachheit und Rechtfertigung des Wissens finden sich in jeweils mehreren der Modelle (Hofer & Pintrich, 1997). Hofer und Pintrich (1997) identifizierten auf Basis der Gemeinsamkeiten zwischen den Modellen vier Kerndimensionen in zwei übergeordneten Bereichen epistemologischer Überzeugungen, die sie für ihr eigenes Modell übernahmen. Dimensionen, die nur in einem der Modelle vorkamen oder die sich nicht auf Wissen sondern auf Lernen, Lehren oder Intelligenz bezogen (*Schnelles Lernen, angeborene Lernfähigkeit* bei Schommer, z.B. *Rolle des Lehrenden* bei Baxter Magolda (1992), wurden dabei bewusst nicht berücksichtigt, da sie nach Hofer nicht in den Kernbereich epistemologischer Überzeugungen fallen, sondern nur in der Peripherie liegen (Hofer & Pintrich, 1997, S. 119). Bei diesen beiden Bereichen, die die Kernstruktur der *individuellen epistemologischen Theorien* bilden, handelt es sich um Überzeugungen zur *Beschaffenheit des Wissens* (*nature of knowlegde*) und zur *Beschaffenheit des Wissensprozesses* (*nature of knowing*) (siehe Abb. 2.2, S. 44). Überzeugungen zur Natur des Wissens betreffen die Art und Weise, wie Wissen selbst beschaffen und organisiert ist („what one believes knowledge is“ Hofer, 2004a, S. 46), während Überzeugungen zur Natur des Wissensprozesses sich darauf beziehen, wie der Prozess des Wissens abläuft und Vorstellungen über die Herkunft und Begründbarkeit von Wissen beinhaltet („how one comes to know“ Hofer, 2004a, S. 46).

| Forschende | Kerndimensionen epistemologischer Überzeugungen: Beschaffenheit des Wissens: Sicherheit | Einfachheit | Beschaffenheit des Wissensprozesses: Quelle | Rechtfertigung | Periphere Dimensionen epistemologischer Überzeugungen: Beschaffenheit des Lernens | Beschaffenheit der Intelligenz |
|---|---|---|---|---|---|---|
| Perry | absolut ↔ kontextabhängiger Relativismus | | Autoritäten ↔ selbst | | | |
| Belenky | | | vermittelt ↔ konstruiert | | | |
| | | | außerhalb des Selbst ↔ Selbst als Bedeutungs-konstrukteur | | | |
| Baxter Magolda | absolut ↔ kontextabhängig | | Verlassen auf Autoritäten ↔ Selbst | Vermittelt oder Können ↔ Evidenz im Kontext beurteilt | Rolle des Lerners, der Peers, der Lehrenden ↔ Evaluation des Lernens | |
| King & Kitchener | sicher, richtig/falsch ↔ unsicher, kontextabhängig | einfach ↔ komplex | Verlassen auf Autoritäten ↔ der Wissende konstruiert selbst Bedeutung | Begründung nicht nötig ↔ Wissen ist konstruiert, Begründungen werden kritisch reevaluiert | | |
| Kuhn | absolut, richtig/falsch Antwort ↔ Wissen nach relativen Vorzügen bewertet | | Experten ↔ Experten kritisch beurteilt | Akzeptanz von Fakten, ungeprüfte Expertise ↔ Bewertung der Expertise | | |
| Schommer | absolut ↔ vorläufig, veränderlich | isolierte, eindeutige Wissensbestandteile ↔ verbundene Konzepte | vermittelt durch Autoritäten ↔ logisch abgeleitet | | schnelles Lernen oder gar nicht ↔ Lernen über Zeit | angeborene Lernfähigkeit ↔ Lernen kann gelernt werden |

ABBILDUNG 2.2: Gemeinsamkeiten in den Dimensionen epistemologischer Überzeugungen zwischen verschiedenen Modellen nach Hofer und Pintrich (1997, S. 113 ff.), in eigener Darstellung

Die beiden Bereiche beinhalten je zwei Kerndimensionen: Überzeugungen zur *Sicherheit des Wissens* und zur *Einfachheit des Wissens* sind der Kategorie *nature of knowledge* zuzuordnen. Überzeugungen zur *Quelle des Wissens* und zur *Begründung des Wissens* gehören dagegen der Kategorie *nature of knowing* an (Hofer & Pintrich, 1997). Die beiden zentralen Bereiche werden seitdem als Kern des Konstrukts epistemologischer Überzeugungen gesehen (Pintrich, 2002) und für die Definition des Konstrukts verwendet (z.B. Greene et al., 2016; Schraw et al., 2017). Es herrscht auch weitgehend Konsens im Forschungsbereich, dass Sicherheit des Wissens, Einfachheit bzw. Struktur des Wissens, Quelle des Wissens und Rechtfertigung bzw. Begründung des Wissens die zentralen Dimensionen darstellen (Bendixen & Rule, 2004, S. 70).

Alle vier Dimensionen werden als Kontinuum mit zwei Polen betrachtet, wobei der eine Pol als niedrigeres Level epistemologischer Überzeugungen und der andere als höheres Level epistemologischer Überzeugungen beschrieben wird. Über diese beiden Pole enthält jede der Dimensionen eine Entwicklungskomponente, ähnlich den Entwicklungsstufenmodellen (vgl. Brownlee, Ferguson & Ryan, 2017). Die folgende Darstellung der Skalen basiert auf Hofer und Pintrich (1997) und Hofer (2000). Die Beschreibung der Dimensionen enthält Verweise auf die Autorinnen und Autoren der von Hofer und Pintrich (1997) rezipierten Modelle, da Hofer und Pintrichs Definitionen der einzelnen Dimensionen auf diese Konzeptualisierungen epistemologischer Überzeugungen zurückgehen.

**Sicherheit des Wissens (*certainty of knowledge*)**

Diese Dimension betrifft im einen Pol die epistemologische Vorstellung, dass Wissen einerseits als absolut angenommen wird, in dem Sinne, dass es mit Sicherheit „die eine richtige Antwort“ –*Wahrheit*– gibt. Damit geht einher, dass eine klare Grenze zwischen *richtigem* Wissen (Wahrheit) und *falschem* Wissen (Unwahrheiten) angenommen wird. Auf der anderen Seite steht die Überzeugung, dass es keine absoluten Wahrheiten gibt und Wissen mit Unsicherheit behaftet ist, d.h. relativ ist, was wahr ist. Individuen in letzterem Extrem gehen davon aus, dass komplexe Probleme mehrere richtige Antworten haben können. Sie sind im Sinne der höchsten Stufe reflektiven Urteilens von King und Kitchener (1994) neuen Interpretationen gegenüber offen und nach dem Vorbild des *Evaluisten* von Kuhn (1991) auch der Möglichkeit, die eigene Ansicht im echten Austausch mit anderen zu verändern.

**Einfachheit bzw. Struktur des Wissens (*simplicity of knowledge*)**

Die Überzeugungen dieser Dimension beziehen sich auf die Struktur und die Organisation des Wissens, die als entweder einfacher oder aber komplexer angesehen wird. Hofer und Pintrich (1997) wählen für die Beschreibung dieser Dimension ganz zentral die Konzeption von Schommer (1990, 1994). Der eine Pol (niedrigeres Level) dieser Dimension betrifft dementsprechend die Überzeugung darüber, dass Wissen aus einer Ansammlung von konkreten, diskreten und erkennbaren Fakten besteht. Im anderen Pol (höheres Level) wird Wissen als untereinander stark verbundene Konzepte und dementsprechend als relativ, kontextabhängig und ungewiss betrachtet.

**Quelle des Wissens (*source of knowledge*)**

Die Beschreibung dieser Dimension leiten Hofer und Pintrich (1997) aus den Gemeinsamkeiten im Verständnis zu dieser Dimension aus den Modellen (Baxter Magolda, 1992; Belenky et al., 1986; King & Kitchener, 1994; Perry, 1970; Schommer, 1990, 1994) ab, die sie in ihrer Arbeit von 1997 untersucht haben (siehe Abb. 2.2, S.44; siehe auch Hofer, 2000, S. 381). Diese Dimension umfasst die epistemologische Vorstellung über den Ursprung des Wissens. Sie bezieht sich auf die Beziehung zwischen der wissenden Person und dem zu wissenden (*knower and known*, Hofer, 2004a, S. 46). Das betrifft im niedrigeren Level die Überzeugung (linker Pol), dass Wissen entweder von außen über Autoritäten an Lernende herangetragen wird oder aber auf höherem Level (rechter Pol) auf der Fähigkeit der Lernenden aufbaut, Wissen in Interaktion mit anderen zu konstruieren.

**Rechtfertigung des Wissens (*justification of knowing*)**

Diese Dimension umfasst, wie Individuen einschätzen, was ein angemessener Wissensanspruch (*knowledge claim*) ist und zwar sowohl in Bezug auf die Bewertung des Wissens als auch die Produktion von Wissen (Weinstock & Cronin, 2003). Sie betrifft, wie Lernende von Belegen Gebrauch machen, Autorität und Expertise annehmen sowie Expertinnen und Experten bewerten. Zur Beschreibung der beiden Pole des Kontinuums dieser Dimension beziehen sich Hofer und Pintrich (1997) auf das *Reflective Judgment* Modell von King und Kitchener (1994). Individuen auf niedrigerem Level (linker Pol) rechtfertigen Wissensansprüche über Beobachtung oder Autorität (wenn Wissen als sicher wahrgenommen wird) oder auf der Basis dessen, was sich für sie richtig anfühlt (wenn Wissen als unsicher wahrgenommen wird). Individuen auf höherem Level (rechter Pol) prüfen Behauptungen, wägen die Ansichten von Experten ab und integrieren sie (King & Kitchener, 1994). Regeln der Überprüfung und Auswertung sowie die Integration verschiedener Wissensquellen dienen dabei als Belege (ebd).

Hofer und Pintrich (1997) betrachten epistemologische Überzeugungen, übereinstimmend mit der Forschung zu *Conceptual Change* (Vosniadou, 1994; Wellman & Gelman, 1992) und *theory of mind* (Wellman, 1990), in ihrer Form und Struktur als persönliche Theorien (*personal theories*) (Hofer & Pintrich, 1997, S. 117). Sie postulieren demnach, dass die individuellen Überzeugungen zu Wissen und dem Wissensprozess zu einem Wissensbereich analog zum Aufbau wissenschaftlicher Theorien in Form von subjektiven Theorien strukturiert ist (Hofer, 2001, S. 360 f.). Dieser Aufbau epistemologischer Überzeugungen stellt nach Hofer und Pintrich (1997) einen geeigneten Mittelweg zwischen den Entwicklungsstufenmodellen und dem *Epistemological Belief System* von Schommer (1990) dar, da Schwierigkeiten beider Modelle hier berücksichtigt würden. Eine Kritik an den Entwicklungsstufenmodellen ist demnach, dass sie ein zwar kohärentes aber zu grobkörniges Erklärungsmodell darstellen: Sie erlauben keine Variationen im Aufbau der epistemologischen Überzeugungen innerhalb einer Stufe, da für jede Stufe genau definiert ist, über welche spezifische Sichtweise

auf Wissen eine Person verfügt, die dieser Stufe zuzuordnen ist. Horizontale Entwicklungsverzögerungen (d.h. nicht vorgesehene Variationen in den individuellen epistemologischen Überzeugungen innerhalb einer Stufe) sind nach diesen Modellen demnach nicht vorgesehen und nicht richtig abbildbar. Schommers Modell betrachtet die Überzeugungen zu Wissen im Gegensatz dazu als orthogonales System von Dimensionen, die sich nicht notwendigerweise zu einer kohärenten Struktur zusammenfügen müssen. Mit dem Erklärungsansatz des Aufbaus epistemologischer Überzeugungen als persönliche Theorien, wird eine gewisse Kohärenz des Überzeugungssystems, d.h. Zusammenhänge der Dimensionen untereinander, angenommen. Gleichzeitig ist das System nicht zu starr und lässt Variationen zu (Hofer & Pintrich, 1997).

Hofer und Pintrich (1997) untersuchen die in ihrem Review diskutierten Modelle epistemologischer Überzeugungen und deren zugrundegelegtes Verständnis des Konstrukts auf ihre Eignung, epistemologische Überzeugungen in ihrem Aufbau als persönliche Theorien zu betrachten. Dazu bedienen sie sich der Kriterien von Wellman (1990). Ein Wissensbestand lässt sich demnach als Theorie bezeichnen, wenn er erstens in seinen Konzepten und Prinzipien über eine gewisse innere Kohärenz verfügt. Die Kohärenz lässt sich nach Wellman (1990) als Kontinuum betrachten, das sich von kaum verbundenen Fakten bis hin zur formalen Theorie erstreckt, die über die höchste Kohärenz verfügt. Zweitens sollte eine Theorie immer deutlich machen, welche Aspekte in den Erklärungsbereich der Theorie fallen und welche nicht. Nicht nur muss der Zielbereich klar abgegrenzt werden, sondern in diesem Bereich muss die Theorie bestimmte Entitäten oder Prozesse unterscheiden und eine Einteilung oder Kategorisierung ermöglichen. Drittens muss ein Wissensbestand, um als Theorie zu gelten, ein kausales Erklärungsmodell der Phänomene in seinem spezifischen Zielbereich liefern. In Bezug auf ihre Kohärenz seien die Überzeugungen einer Person über das eigene Wissen und den Wissensprozess (sofern es sich bei dieser Person nicht gerade um einen Philosophen handele) natürlich nicht mit einer formalen, wissenschaftlichen Theorie gleichzusetzen. Sie ließen sich jedoch als subjektive Theorie beschreiben, die dem Individuum dazu dienen, sich im Alltag ihre Umwelt zu strukturieren, zu erklären und vorherzusagen (Hofer & Pintrich, 1997). Das zweite Kriterium wird nach Hofer und Pintrich (1997) insofern erfüllt, als dass in den verschiedenen Modellen zu epistemologischen Überzeugungen übereinstimmende Aussagen darüber getroffen werden, was zum Konstrukt gehört und was nicht. Einerseits konnte eine modellübergreifende Einteilung in *Überzeugungen zum Wissen* und *Überzeugungen zum Prozess des Wissens* identifiziert werden. Darüber hinaus wurden nicht nur übereinstimmend Überzeugungen zu den Kernbereichen Sicherheit, Quelle, Einfachheit und Rechtfertigung des Wissens identifiziert, sondern es konnte im Rahmen der Forschung zu den verschiedenen Modellen auch schon gezeigt werden, dass diese Überzeugungen, wie subjektive Theorien, die individuelle Wahrnehmung von Erfahrungen beeinflussen (Hofer & Pintrich, 1997, S. 118). Hofer und Pintrich (1997) sind in ihrem Review noch sehr vorsichtig in ihren Aussagen darüber, ob ihr Modell auch das dritte von Wellmans Kriterien zu erfüllen vermag. Sie sehen allerdings schon erste Hinweise auf mögliche kausale Erklärungsmodelle, z.B. insofern, als dass spezifische, individuell vorliegende Überzeugungen zum Wissen sich einschränkend auf den Prozess des Wissens auswirken

mögen (Hofer & Pintrich, 1997, S. 118). Später wird auf Basis von Erkenntnissen zur Disziplinspezifität (Donald, 1995; Hofer, 2000) epistemologischer Überzeugungen deutlich, dass ihr Modell epistemologischer Theorie insofern schon das dritte Kriterium erfüllt, als dass es u.a. ein kausales Erklärungsmodell für vorliegende intraindividuelle disziplinspezifische Unterschiede in den epistemologischen Überzeugungen zu liefern vermag (Hofer, 2001). Hofer und Pintrich nehmen also an, dass die vier identifizierten Kerndimensionen epistemologischer Überzeugungen „are related to each other in coherent and internally consistent ways, do make some important distinctions about knowledge, and may provide a causal-explanatory framework for thinking about knowledge." (Hofer & Pintrich, 1997, S. 119). Insofern vereinen die *Epistemological Theories* (Hofer, 2001; Hofer & Pintrich, 1997) die in den Entwicklungsmodellen angenommene kohärente Struktur epistemologischer Überzeugungen, indem sie von einer höheren Integration zwischen den individuellen Überzeugungsbereichen ausgehen, mit der bei Schommer angenommenen expliziten Mehrdimensionalität (Hofer, 2001) und liefern darüber hinaus ein Erklärungsmodell für Unterschiede in epistemologischen Überzeugungen in Abhängigkeit von Fachdisziplinen.

Damit steht diese Konzeption, im Gegensatz zu den früheren Modellen, explizit im Einklang mit der heute anerkannten Annahme (vgl. z.B. Hofer, 2016), dass individuelle epistemologische Überzeugungen von der jeweiligen Fachdisziplin abhängig sind und dieselbe Person deshalb sowohl über allgemeine und gleichzeitig auch über disziplinspezifische epistemologische Überzeugungen verfügen kann (Hofer, 2001, S. 362). Hervorzuheben ist, dass dies zu jener Zeit einer der wenigen ersten Vorstöße in die Erforschung (Donald, 1995; Hofer, 2000) der Disziplinabhängigkeit epistemologischer Überzeugungen darstellte.

Um vermutete disziplinspezifische Unterschiede feststellen zu können und ganz allgemein die Dimensionalität epistemologischer Überzeugungen zu prüfen, entwickelte Hofer (2000) den englischsprachigen *Disciplin-Focused Epistemological Belief Questionnaire* (DEBQ). Hofer und Pintrich (1997) nahm zunächst die auf Basis ihres Reviews identifizierten vier Dimensionen an, die oben ab Seite 45 beschrieben wurden. Auf Basis von Perrys CLEV und Schommers SEQ und in Übereinstimmung mit der Literatur des Reviews entwickelte Hofer (2000) 27 fünfstufige Likert-Items (1 =*strongly disagree* bis 5 = *strongly agree*). Die spezifische Disziplin, auf die sich die jeweilige Erhebung beziehen soll, muss den Versuchspersonen über die Instruktion mitgeteilt werden. In den Items wird über Formulierungen wie „in this field . . . " ein Bezug zur betreffenden Disziplin hergestellt.

Der Fragebogen wurde von Hofer (2000) bei einer Stichprobe von 326 Studierenden im ersten Studienjahr im Rahmen einer einführenden Psychologieveranstaltung eingesetzt. Die Studierenden füllten den DEBQ je einmal für Psychologie und Naturwissenschaft aus. Zur Validierung des DEBQ setzte Hofer zusätzlich eine gekürzte Version des SEQ-M von Schommer nach Qian und Alvermann (1995) ein. Über die faktorenanalytische Auswertung identifizierte Hofer eine andere Faktorenstruktur als ursprünglich angenommen. Zunächst fand Hofer einen neuen, unerwarteten Faktor, der sich auf die Dimension *attainability of truth* (Beispielitem: „Experts in this field can ultimately get to the truth.") bezieht. Die Items zu den angenommenen Skalen *Sicherheit* und *Einfachheit* schlossen sich zu einer gemeinsamen Skala

zusammen: *Certainty/simplicity of knowledge* (Beispielitem: „Most of what is true in this subject is already known."). Die ursprünglich angenommenen Dimensionen *Begründung* und *Quelle* des Wissens wurden durch die identifizierten Faktoren nicht in voller inhaltlicher Breite abgedeckt, sondern bezogen sich je auf einen Teilaspekt: *justification of knowing: personal* (Beispielitem: „First hand experience is the best way of knowing something in this field.") und *source of knowledge: authority* (Beispielitem: „If you read something in a textbook for this subject, you can be sure that it is true." Hofer, 2000). Hofer (2000) konnte intraindividuelle disziplinspezifische Unterschiede zwischen Psychologie und Naturwissenschaft zeigen. Die Ergebnisse ließen sie annehmen, dass neben den disziplinspezifischen auch übergreifende, allgemeine epistemologische Überzeugungen wirken.

Zwar gelingt es Hofer (2000) im Rahmen ihrer Instrumentenentwicklung nicht, die vier ursprünglich angenommenen Dimensionen faktorenanalytisch abzubilden, dies lässt jedoch keinen Zweifel daran aufkommen, dass es sich bei diesen vier zentralen Dimensionen um die Kerndimensionen epistemologischer Überzeugungen handelt. Schließlich bilden die Dimensionen Quelle, Rechtfertigung, Sicherheit und Einfachheit nach wie vor die Schnittmenge der verschiedensten theoretischen Konzeptionen epistemologischer Überzeugungen und konnten in anderen Erhebungskontexten schon vielfach abgebildet werden.

Der bedeutende Beitrag zur Forschung Hofer und Pintrichs (1997) besteht in der Identifizierung der beiden übergreifenden Bereiche und vier Kerndimensionen epistemologischer Überzeugungen, die sich bis heute in den verschiedenen modernen Konzeptionen wiederfinden. Darüber hinaus nahmen sie schon damals an, dass die epistemologischen Überzeugungen einer Person von der sie betreffenden Fachdisziplin abhängig sein müssten, was sich später (Hofer, 2000) bestätigte. Sie entwickelte daraufhin mit dem DSBQ (Hofer, 2000) einen Fragebogen, der diese Disziplinspezifität berücksichtigt.

Hofer ist seit ihren ersten Forschungsbeiträgen über die Jahre hinweg bis heute forschend aktiv und befasste sich u.a. auch mit dem Einfluss epistemologischer Überzeugungen auf den Lernprozess (Hofer, 2001) und den Beziehungen zur Metakognition. Ihr Modell zum Einfluss epistemologischer Überzeugungen auf den Lernprozess wird im Rahmen des betreffenden Kapitels (siehe Kapitel 2.8.1, siehe Abb. 2.8, S. 87) dargestellt. In einer späteren Weiterentwicklung der Konzeption verortet Hofer (2004a) epistemologische Überzeugungen in Bezug zu Kognition und Metakognition. Dies wird im Zusammenhang mit der Forschung zur Verortung epistemologischer Überzeugungen in den Theorien der Kognition und Metakognition (siehe Kapitel 2.7, S. 80) dargestellt.

Unter Berücksichtigung der oben beschriebenen Dimensionen sowie neuerer Forschungsbeiträge von Hofer lässt sich ihre Definition epistemologischer Überzeugungen folgendermaßen zusammenfassen: Sie betrachtet epistemologische Überzeugungen als System aus in Theorien organisierten Überzeugungsdimensionen, die Wissen und den Wissensprozess betreffen (Hofer, 2005). Die Dimensionen sind Überzeugungen zur Quelle des Wissens, Rechtfertigung/Begründung des Wissensprozesses, Sicherheit des Wissens und Einfachheit des Wissens (Hofer & Pintrich, 1997). Diese individuellen Überzeugungen entwickeln sich über

die Zeit in eine einigermaßen vorhersagbare Richtung weiter (Hofer, 2005). Epistemologische Überzeugungen werden im Kontext aktiviert, d.h. werden durch den Kontext (z.B. instruktionaler Kontext, akademischer Kontext, sozio-kultureller Kontext) auch beeinflusst (vgl. Hofer, 2005, 2006a). Sie arbeiten sowohl kognitiv als auch metakognitiv (vgl. Hofer, 2004a, 2005). Darüber hinaus nimmt sie drei Ebenen epistemologischer Überzeugungen an: allgemeine sowie disziplinspezifische Überzeugungen zu Wissen und dem Wissensprozess sowie die Perspektive der Disziplin auf Wissen und den Wissensprozess (vgl. Hofer, 2005, 2006b).

Die zuvor dargestellten zentralen Konzeptionen epistemologischer Überzeugungen von Perry (1970), Kuhn (1991), King und Kitchener (1994) sowie Schommer (1990) stellen unterschiedliche Definitionen des Konstrukts dar, indem sie das Konstrukt abgrenzen und aufzeigen, was dieses umfasst und welche Struktur es aufweist. Wie Hofer und Pintrich (1997) in ihrem Review jedoch zeigen konnten, weisen diese Konzeptionen Gemeinsamkeiten auf, die sich als Kerndimensionen epistemologischer Überzeugungen beschreiben lassen und aus denen Hofer und Pintrich (1997) ihr Modell entwickelten.

## 2.3 Entwicklung und Veränderung epistemologischer Überzeugungen

Es gibt einige quantitative und qualitative Studien (z.B. Burr & Hofer, 2002; Conley et al., 2004; Kuhn & Pearsall, 2000), die zeigen können, dass Überzeugungen zu Wissen und dem Wissensprozess sich – durch kognitive Entwicklungen und Bildungsentwicklungen verursacht – über die Zeit verändern (Grossnickle, Alexander & List, 2017). In Bezug auf diese Entwicklung epistemologischer Überzeugungen stellen sich drei zentrale Fragen. Die erste Frage betrifft das Problem, *was* sich bezüglich des epistemologischen Verständisses von Wissen denn genau entwickelt (Conley et al., 2004). Sie wurde sowohl für die entwicklungspsychologische als auch für die pädagogisch-psychologische Sichtweise oben im Rahmen der Darstellung der verschiedenen Modelle bereits beantwortet (siehe Kapitel 2.2 ab S. 25). Hofer und Pintrich (1997) beispielweise, definieren über ihre Dimensionen *Quelle des Wissens*, *Einfachheit des Wissens*, *Sicherheit des Wissens* und *Begründung des Wissens*, *was* sich über die Zeit weiterentwickeln kann. Die zweite zentrale Frage betrifft daran anschließend die Beschaffenheit der Entwicklung (Conley et al., 2004), während die dritte Frage das *Wie* betrifft. Wie entwickeln sich epistemologische Überzeugungen weiter? Durch was wird die Entwicklung ausgelöst? Wie kann die Entwicklung beeinflusst werden?

### 2.3.1 Beschaffenheit der Entwicklung

In Bezug auf die Beschaffenheit der Entwicklung wird mit großer Übereinstimmung von einer individuellen Entwicklung von weniger sophistizierten, naiven bis hin zu sophistizierten Überzeugungen über Wissen und den Wissensprozess ausgegangen (z.B. Greene et al., 2010; Hofer & Pintrich, 1997; King & Kitchener, 2004; Perry, 1970), die u.a. durch akademische Erfahrungen verursacht wird (Greene et al., 2010; Hofer & Pintrich, 1997). Die Entwicklungsstufenmodelle gehen übereinstimmend von einer generellen Entwicklung vom Absolutismus über den Relativismus bis hin zum Evaluatismus aus, unterscheiden sich allerdings in ihrer Zahl der Abstufung (von 3 bis 9) (Hofer, 2005). Letztgenannte gilt als die sophistizierteste Entwicklungsstufe. Hofer und Pintrich (1997) beantworten die Frage nach der Natur der Entwicklung, indem sie die entwicklungspsychologische Sichtweise integrieren. Sie gehen davon aus, dass Individuen sich innerhalb jeder der vier Dimensionen, die je ein Kontinuum darstellen, über die Zeit in Richtung einer sophistizierteren Ausprägung entwickeln (vgl. Beschreibung der Dimensionen auf Seite 45). Allerdings herrscht Uneinigkeit darüber, wie und wann die Veränderungen stattfinden (Bendixen & Rule, 2004). Auffällig ist auch, dass sich die Untersuchungszeiträume in Bezug auf die Natur der Entwicklung epistemologischer Überzeugungen massiv unterscheiden. Die explizit entwicklungspsychologisch ausgerichtete Forschung untersuchte die Entwicklung größtenteils längsschnittlich über sehr lange Zeiträume hinweg (ebd.). Perry (1970) erforschte beispielsweise die epistemologische Entwicklung von College Studierenden über einen Zeitraum von vier Jahren. Kuhn (1991) verglich dagegen mittels eines querschnittlichen Designs die epistemologischen Überzeugungen von Individuen unterschiedlichen Alters. Hallett, Chandler und

Krettenauer (2002) dagegen verglichen querschnittlich die epistemologischen Überzeugungen von Highschool Schülerinnen und Schülern, Studierenden im ersten und zweiten Studienjahr und solchen im dritten und vierten Semester. Auch Jehng et al. (1993) und Schommer (1993a) fanden in ihren Studien Unterschiede in den Ausprägungen der epistemologischen Überzeugungsdimensionen zwischen Studierenden niedriger und höherer Semester. Die Studierenden im fortgeschrittenen Studium verfügten dabei auch über die fortgeschritteneren epistemologischen Überzeugungen. Conley et al. (2004) konnten zeigen, dass sich die epistemologischen Überzeugungen von Schülerinnen und Schülern der fünften Jahrgangsstufe über einen Zeitraum von neun Wochen weiterentwickelten.

Abbildung 2.3 von Wildenger et al. (2010, S. 227) zeigt die zeitliche Entwicklung epistemologischer Überzeugungen von der Kindheit bis ins junge Erwachsenenalter, die auf der Basis verschiedener Forschungsergebnisse mit unterschiedlichen zugrundeliegenden theoretischen Konzeptionen zusammengestellt wurde. Die Angaben in Klammern sind nach Wildenger et al. (2010, S. 227) nicht hinreichend empirisch belegt (siehe Abb. 2.3). Hier zeigt sich, dass sich die Entwicklung vom Absolutismus – auch Dualismus genannt– über den Relativismus bis zum Evaluatismus in verschiedenen Modellen wiederfindet.

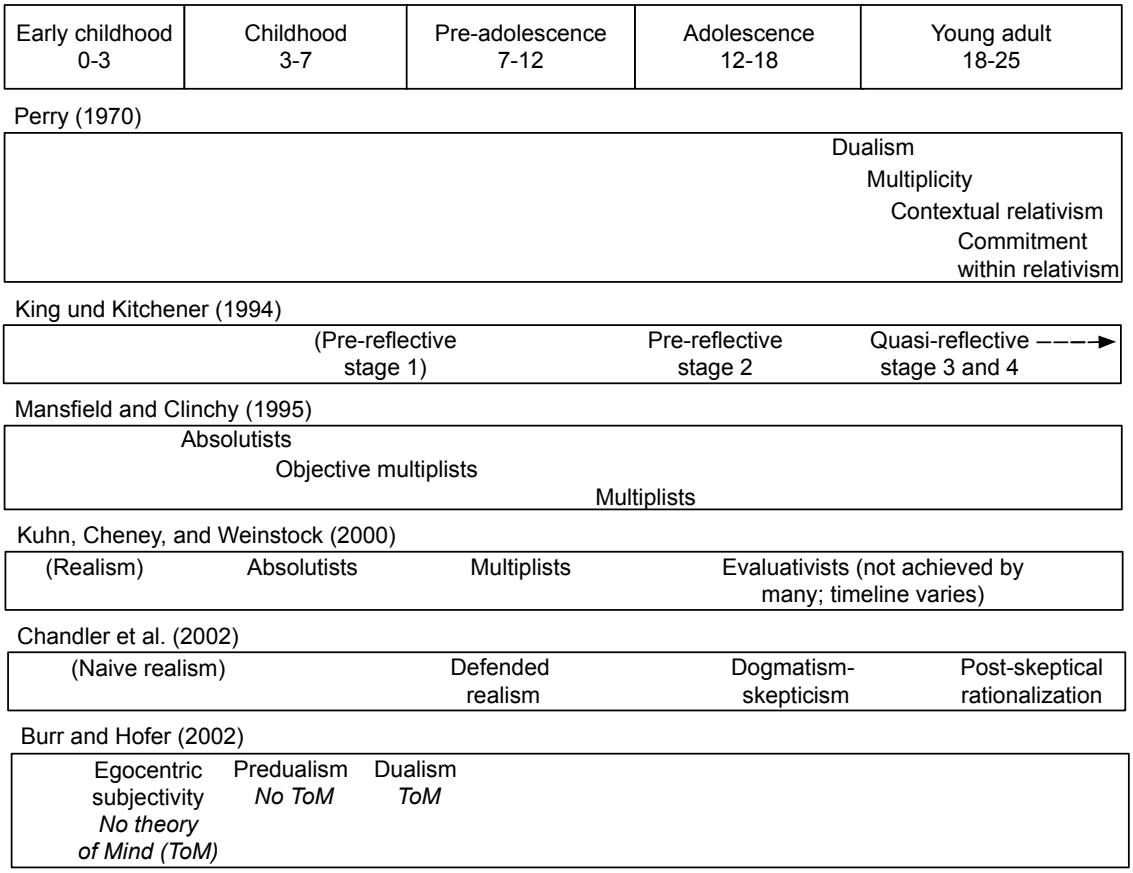


ABBILDUNG 2.3: Zeitliche Entwicklung der epistemologischen Überzeugungen in verschiedenen Modellen von Wildenger et al. (2010, S. 227), in eigener Darstellung

Das Alter, in dem Personen die entsprechenden Entwicklungsstufen erreichen, unterscheidet sich je nach Modell beträchtlich (siehe Abb. 2.3, S. 52). Die Entwicklung über die Stufen hinweg ist bei Perry (1970) vollständig in der Adoleszenz und im jungen Erwachsenenalter verortet. In Perrys Studie erreichten mit jedem aufeinanderfolgenden Jahr (vom ersten College Semester bis zum letzten) mehr College Studierende höhere Stufen des Denkens in den durchgeführten Interviews. Baxter Magolda (2002) kam mit einer neueren längsschnittlichen Studie bei College Studierenden zu vergleichbaren Ergebnissen. Der dominate Wissensmodus *transitional knowing* wurde in ihrer Studie von 32% der Studierenden im ersten Studienjahr, 53% im zweiten, 83% im dritten, und 80% im vierten Studienjahr eingesetzt (Baxter Magolda, 2002). Moore (1991) fand beim Einsatz des *Measure of Intellectual Development* (MID), ein Erhebungsinstrument zum *Perry Scheme*, jedoch nur sehr wenig Veränderung bei 18 bis 21 Jährigen.

Im Gegensatz dazu postulieren andere Modelle eine viel frühere Entwicklung epistemologischer Überzeugungen (siehe Abb. 2.3, S. 52). King, Kitchener und Kollegen (King, 1977; King & Kitchener, 1994, 2002; King et al., 1983; Kitchener & King, 1981; Kitchener, King, Wood & Davison, 1989; Kitchener, Lynch, Fischer & Wood, 1993; Kitchener & Wood, 1987) erforschten epistemologische Überzeugungen im Gegensatz zu Perry bei vielen verschiedenen Altersgruppen, von Highschool Studierenden bis Personen im späten Erwachsenenalter. Sie gehen jedoch davon aus, dass die epistemologische Entwicklung einerseits viel früher beginnt aber andererseits auch nur sehr wenige Personen (z.B. solche mit PhD in Philosophie) die höchste Entwicklungsstufe erreichen (Hallett et al., 2002, S. 288). Nach den Ergebnissen der Studien von King und Kitchener (1994) mit über 1700 Personen erreichten College Schülerinnen und Schüler im ersten und zweiten Jahr durchschnittlich Stufe 3.6, im dritten Jahr Stufe 3.7 und im vierten Jahr Stufe 4.0. Erwachsene College Schülerinnen und Schüler unterschieden sich dabei in ihrer Entwicklungsstufe nicht von den College Schülerinnen und Schülern im typischen Alter (Adoleszenz). College-Absolventen befanden sich durchschnittlich auf der 4. Stufe, d.h. auf der niedrigsten Stufe quasi-reflektierten Schlussfolgerns. Bei Hochschulabsolventen (*graduate students*), die sich im Doktorandenprogramm (*graduate training*) befanden, war der stärkste Entwicklungszuwachs zu beobachten. In der Anfangsphase des Doktorandenprogramms erreichten die Befragten durchschnittlich 4.6 und in den späteren Phasen des Trainings durchschnittlich 5.3. Konsistent reflektiertes Schlussfolgern und Problemlösen auf Stufe 6, erreichten nur Personen auf fortgeschrittenem Doktorandenniveau (*advanced doctoral level*). Hochschulabsolventen (*graduate students*) in den Sozialwissenschaften erreichten durchschnittlich eine höhere Entwicklung, als solche der mathematischen Wissenschaften. Als Ursache dieses Unterschieds wird das vermehrte Arbeiten mit komplexen, schlecht strukturierten Problemen in den Sozialwissenschaften vermutet. Studien mit Erwachsenen, bei denen es sich nicht um Studierende handelte, zeigten, dass der vorhergehende Collegebesuch sehr stark mit dem Level des reflektierten Denkens in Verbindung steht (Hofer & Pintrich, 1997).

Haerle (2006) konnte für den Bereich der Naturwissenschaften zeigen, dass bereits Grundschulkinder über Überzeugungen zum Wissen und Wissensprozess verfügen. Weitere Befunde bestätigen dies für die Naturwissenschaften (z.B. Elder, 2002; Thoermer & Sodian,

2002). Kuhn (1991) untersuchte die epistemologischen Überzeugungen bei 169 Schülerinnen und Schülern (im 6., 9. und 12. Schuljahr) sowie bei Erwachsenen in ihren 20ern, 30ern und 60ern. Die jüngsten Befragten erreichten generell die niedrigsten drei Stufen (0, 1 und 2), während nur die Ältesten und am besten ausgebildeten Individuen die höchste, 6. Stufe (*evaluatives Schlussfolgern*) erreichten.

Der epistemologischen Entwicklung von Teenagern widmeten sich Chandler und Kollegen (z.B. Boyes & Chandler, 1992; Chandler, 1975, 1987; Chandler, Boyes & Ball, 1990). Sie konnten bei manchen Achtklässlern bereits relativistische Überzeugungen zu Wissen und dem Wissensprozess identifizieren. In der zwölften Jahrgangsstufe zeigten viele bereits Überzeugungen, die den späteren *commitment*-Stufen nach Perry zuzuordnen sind (Boyes & Chandler, 1992). Relativistisches Verstehen, d.h. bereits fortgeschrittenere epistemologische Überzeugungen, die sich nach Perry erst im College-Alter entwickeln, konnten vielfach schon vor dem jungen Erwachsenenalter und sogar noch vor der Adoleszenz identifiziert werden (Reich, Oser & Valentin, 1994; Smith, Maclin, Houghton & Hennessey, 2000; Walton, 2000). Im Zusammenhang mit anderen, möglicherweise verwandten Theorien der kognitiven Entwicklung, der *Theory-of-Mind* Forschung (Hofer, 2005, S. 99) konnte gezeigt werden, dass selbst Vierjährige schon erste Einblicke zur subjektiven Beschaffenheit des Wissens haben (z.B. Gopnik & Wellman, 1992; Perner & Davies, 1991; Wellman, 1990), während Chandler und Kollegen ein erstes Verstehen über die Personenabhängigkeit von Wissen, das sie *interpretive theory-of-mind* nennen, im Alter von sieben bis acht verorten.

Auch Hofer (2005) sieht die Anfänge epistemologischer Überzeugungen im Alter von drei bis fünf und stellt eine Verbindung zur *Theory-of-Mind* her. Epistemologisches Verstehen scheint bei Kindern zwischen drei und fünf Jahren mit einer egozentrischen Subjektivität (*egocentric subjectivity*) zu beginnen (Burr & Hofer, 2002). Diese epistemologische Sichweise auf Wissen kann sich erst verändern, wenn das Kind versteht, dass andere Personen unterschiedliche Überzeugungen, Intentionen und Wünsche haben können (Hofer, 2005, S. 99), d.h. durch die Entwicklung der *Theory-of-Mind* (Wellman, 1990; Wellman, Cross & Watson, 2001). Dies leitet die absolutistische Sichweise auf Wissen ein, in der Wissen als objektiv und sicher angesehen wird. Wenn eine Person die potenzielle Unsicherheit von Wissen erkennt, wird die absolutistische Sichtweise von der relativistischen abgelöst. Die Person kehrt wieder zur subjektiven Sichtweise auf Wissen zurück, weiß nun allerdings, dass verschiedene Personen unterschiedliche subjektive Meinungen zu etwas haben können. Das Individuum verfügt allerdings noch nicht über die Fähigkeit, zwischen den unterschiedlichen Sichtweisen anderer Personen zu differenzieren, d.h. diese unterschiedlich zu gewichten. Wenn die richtigen Voraussetzungen der Entwicklung gegeben sind, ist es dem Individuum in einem weiteren Entwicklungsschritt möglich, die Objektivität mit der Subjektivität von Wissen zu vereinbaren. Das Individuum lernt mit der Unsicherheit des Wissens umzugehen, Wissensansprüche einzuschätzen und sich für Standpunkte zu entscheiden, d.h. diese höher zu gewichten (Hofer, 2005, S. 99). Aus dieser Beschreibung des Entwicklungsverlaufs geht hervor, dass die zentrale Entwicklungsaufgabe epistemologischen Verstehens für Individuen darin besteht, Subjektivität und Objektivität des Wissens zu koordinieren (Burr & Hofer, 2002; Hofer, 2005; Kuhn & Pearsall, 2000; Mansfield & Clinchy, 2002).

Der globale Entwicklungstrend von Dualismus (auch Absolutismus) zum Relativismus (auch Multiplizismus) bis zum Evaluatismus scheint einerseits geklärt (Hofer, 2005). Andererseits werden die Anfänge der Entwicklung epistemologischer Überzeugungen je nach zugrundeliegendem Modell in einem anderen Alter gesehen, während aber die gleichen Stufen und Entwicklungsverläufe für die unterschiedlichsten Altersgruppen identifiziert wurden (ebd.). Diese Befunde (siehe auch Abb. 2.3, S. 52) erscheinen widersprüchlich. Die Ursache dafür, dass es z.B. Perry nicht gelang, epistemologische Überzeugungen schon bei Jugendlichen zu identifizieren, könnte ein Artefakt der eingesetzten Erhebungsmethoden sein (Hallett et al., 2002, S. 288), die beispielsweise aufgrund des Themas oder der Instruktion für bestimmte Altersgruppen ungeeignet sein mögen. Dieses Phänomen kann noch umfassender darüber erklärt werden, dass die epistemologische Entwicklung nicht nur in eine Richtung verlaufen kann, sondern eine Rekursion in frühere Stufen möglich ist (Boyes & Chandler, 1992), die Entwicklung also nicht linear, sondern eher spiralförmig verläuft (Chandler et al., 2002). Diese Ansicht spiegelt sich auch in aktuelleren Modellen wider, wie beispielsweise dem *integrative personal epistemology model* von Bendixen und Rule (2004) und dem TIDE Modell von Muis et al. (2006).

King und Kitchener (1994, 2004) gehen übereinstimmend mit der *complex stage theory* von Rest (1979) davon aus, dass gar nicht davon gesprochen werden kann, dass sich Personen zu einem bestimmten Zeitpunkt in einer bestimmten Stufe befinden. Die Entwicklung des reflektierten Schlussfolgerns charakterisieren King und Kitchener (1994, 2004) als „[...] waves across a mixture of stages, where the peak of a wave is the most commonly used set of assumptions. While there is still an observable pattern to the movement between stages, this development is better described as the changing shape of the wave rather than as a pattern of uniform steps interspersed with plateaus."(siehe King & Kitchener, 1994, S. 140). Die Entwicklung des reflektierten Schlussfolgerns kann nach dieser Ansicht also eher als Welle beschrieben werden, die eine Mischung der epistemologischen Ansichten verschiedener Stufen umfasst. Die Spitze der Welle stellt dabei die hauptsächlich genutzen epistemologischen Annahmen einer bestimmten Stufe dar, während die Ausläufer der Welle die weniger genutzten epistemologischen Annahmen anderer Stufen repräsentieren. Auch wenn ein generelles Fortschreiten durch die Stufen beobachtet werden kann, lässt sich die Entwicklung des reflektierten Schlussfolgerns einer Person eher als eine Veränderung der Wellenform beschreiben. Das heißt, es verändert sich über Zeit, welche der epistemologischen Stufen die Spitze der Welle und welche die Ausläufer darstellen und somit welche die meistgenutzten und weniger genutzten epistemologischen Annahmen der Person bilden (vgl. auch King & Kitchener, 2004, S. 10). Eine Person verfügt nach King und Kitchener (2004, S. 9) zu jeder Zeit über ein spezielles *epistemological framework*, eine spezielle Zusammenstellung epistemologischer Überzeugungen, die ihr zur Interpretation ihrer Erfahrungen dient. Dieses individuelle *epistemological framework* entwickelt sich über die Zeit, insofern als dass es komplexer, integrierter und vollständiger wird.

Nach dem Modell von Baxter Magolda (2004) zur epistemischen Reflexion ist die persönliche Epistemologie sozial konstruiert und kontextabhängig. Die Veränderung ist eher als ein Wandel in Richtung eines komplexeren Systems von epistemologischen Annahmen

zu betrachten, denn als Aneignung spezifischer Lernstrategien oder Fähigkeiten (Bendixen & Rule, 2004).

Im Bezug auf die Natur der Entwicklung wurde oben bereits erwähnt, dass ein evaluatistisches Verständnis von Wissen in diesem Forschungsfeld als höchste, reifste, fortgeschrittenste Form epistemologischer Überzeugungen betrachtet wird. Dieses wird allerdings nach verbreiteter Ansicht (z.B. Hallett et al., 2002; King & Kitchener, 1994; Kuhn & Weinstock, 2002) nur von sehr wenigen Menschen erreicht.

Wie in diesem Abschnitt gezeigt wurde, entwickeln sich die individuellen Überzeugungen zum Wissen und Wissensprozess im Verlauf der Kindheit und bis ins Erwachsenenalter und werden über die Zeit sophistizierter. Im folgenden Abschnitt wird zunächst der Mechanismus der Veränderung dargestellt, bevor auf Auslöser der epistemologischen Entwicklung und Veränderung eingegangen wird.

### 2.3.2 Auslöser von Entwicklung und Veränderung

Als einer der ersten Auslöser und Voraussetzung der Entwicklung epistemologischer Überzeugungen im Leben eines Individuums wird die *Theory-of-Mind* genannt (Burr & Hofer, 2002). *Theory-of-Mind* als Teil des kognitiven Entwicklungsprozesses bezeichnet die kognitive Fähigkeit, zwischen den eigenen mentalen Zuständen (z.B. Wünschen, Überzeugungen, Intentionen) und denen anderer Personen zu unterscheiden (Sodian & Kristen, 2016). Neben dieser grundlegenden Voraussetzung für die Entwicklung epistemologischer Überzeugungen konnte gezeigt werden, dass die Entwicklung und Veränderung der individuellen Überzeugungen zu Wissen und dem Wissensprozess durch akademische Erfahrungen (Greene et al., 2010; Hofer & Pintrich, 1997), Bildung (King & Kitchener, 1994), Erziehung und den kulturellen Hintergrund verursacht wird (vgl. z.B. Kuhn, 1991; Dann 1994; Müller 2009; Mokwinski 2010; Blignaut et al. 2010; Conley et al., 2004; Kuhn, Cheney und Weinstock, 2000). Bildung beeinflusst die epistemologische Entwicklung, indem sie die Fertigkeiten von Individuen stärkt, Informationen kritisch zu beurteilen sowie miteinander in Konflikt stehende Annahmen, Theorien und Evidenzen zu koordinieren (King und Kitchener, 1994, Kuhn 1991). In diesem Zusammenhang bleibt aber noch die Frage nach dem der Veränderung zugrundeliegenden Mechanismus offen, das heißt die Frage danach, wie Bildung, akademische Erfahrungen, Erziehung und der kulturelle Hintergund denn genau zur epistemologischen Entwicklung führen und wie dieser Prozess abläuft.

Viele der Konzeptualisierungen epistemologischer Überzeugungen greifen explizit oder implizit auf Piagets Arbeiten zurück, um die Veränderung epistemologischer Überzeugungen zu erklären. Zentral in Piagets (1950, 1970, 1985) Forschung ist, wie sich individuelle Konzeptionen zu Wissen und dem Wissensprozess über den Prozess der Äquilibration entwickeln (vgl. Rule & Bendixen, 2010, S. 96). Der Prozess der Äquilibration führt über Akkomodation und Assimilation neuer Informationen zur Veränderung des kognitiven Systems über die Zeit. Akkomodation bedeutet die Anpassung der individuellen kognitiven Wissensstrukturen aufgrund von neuen Informationen, während Assimilation die Integration

neuer Informationen in existierende kognitive Strukturen bezeichnet (vgl. Rule & Bendixen, 2010, S. 110 f.). Kognitives Ungleichgewicht wird vielfach (z.B. Baxter Magolda, 2004; King & Kitchener, 2004) als treibende Kraft der Entwicklung und Veränderung epistemologischer Überzeugungen angesehen (Bendixen & Rule, 2004, S. 70). Auch Modelle wie die von King und Kitchener (1994) und Perry (1970) werden, was die Aneignung epistemologischer Überzeugungen und deren Veränderung über Zeit angeht, als konsistent mit dem Piagetschen Konzept des Äquilibriums bezüglich der kognitiven Entwicklung betrachtet (Hofer & Pintrich, 1997). King und Kitchener (1994) postulieren beispielsweise, dass das reflektierte Schlussfolgern einer Person sich weiterentwickelt, wenn die Überzeugungen zu Wissen, über die diese Person gerade verfügt während sie handelt, nicht mit den Erfahrungen vereinbar sind, die die Person in der Umwelt macht. Auch die Konzeptionen von Hofer (2004a), Schommer-Aikins (2004) und Bendixen und Rule (2004) sind mit Annahme des kognitiven Ungleichgewichts als Veränderungsmechanismus vereinbar (Bendixen & Rule, 2004, S. 70). In Modellen, die dem Kontext eine bedeutende Rolle einräumen (z.B. Louca et al., 2004), wird davon ausgegangen, dass die kognitive Dissonanz bei der intentionalen Veränderung epistemologischer Überzeugungen eine Rolle spielen könnte, aber für deren Entstehung bzw. deren allgemeine Entwicklung nicht nötig ist (Pintrich, 2002, Bendixen & Rule, 2004, S. 70 f.). Louca et al. (2004) dagegen sehen die Entwicklung epistemologischer Überzeugungen als eher assoziiert mit den verschiedenen, sich verändernden Kontexten, statt mit der kognitiven Entwicklungsstufe eines Individuums (siehe hierzu auch Hammer & Elby, 2002, Bendixen & Rule, 2004, S. 70). Auch wenn die oben genannten Konzeptionen alle dem kognitiven Ungleichgewicht eine Rolle bei der Veränderung epistemologischer Überzeugungen zusprechen, wird der genaue Mechanismus der Veränderung in diesen nicht erklärt (vgl. Bendixen & Rule, 2004; Rule & Bendixen, 2010).

Die Prozesse, die dazu führen, dass kognitives Ungleichgewicht zur Weiterentwicklung der epistemologischen Überzeugungen führt, beschreiben (Bendixen & Rule, 2004) im Rahmen ihres *Integrative Model for Personal Epistemology* (IM). Dieses wird in Abbildung 2.4 auf Seite 58 dargestellt. Sie beschreiben im Rahmen des IM den Mechanismus, der die individuelle Entwicklung der epistemologischen Überzeugungen verursacht und wie dieser Entwicklungsprozess abläuft (Bendixen & Rule, 2004; Rule & Bendixen, 2010). Dabei stützen sie sich theoretisch stark auf Piagets (1950; 1985) Äquilibrationsprozess als kausalem Veränderungsmechanismus der individuellen kognitiven Entwicklung (vgl. Rule & Bendixen, 2010, S. 110 ff.) sowie auf die Forschung zum *Conceptual Change* und das betreffende Modell von Dole und Sinatra (1998). Bendixen und Rules (2004) Modell ist mit den vorher vorgestellten Entwicklungsstufenmodellen nicht vergleichbar. Einerseits wird hier explizit von verschiedenen Dimensionen epistemologischer Überzeugungen ausgegangen und andererseits handelt es sich bei ihrem Veränderungsprozess um einen dynamischen Prozess, der von vielen Faktoren, wie dem Kontext, Emotionen, Umfeld und Metakognition beeinflusst wird (ebd.). Bendixen und Rule (2004) betrachten epistemologische Überzeugungen übereinstimmend mit der Konzeption von Hofer und Pintrich (1997), Hofer (2000) sowie Pintrich (2002) als

mehrdimensionales Konstrukt, das die Dimensionen *Überzeugungen zur Sicherheit des Wissens*, *Einfachheit des Wissens*, *Rechtfertigung des Wissensprozesses* und *Quelle des Wissens* umfasst. Diese werden in Abbildung 2.4 als Rauten (CK, JK, SoK, SK) im Kasten *current beliefs* dargestellt. Im Rahmen ihres Modells (siehe Abb. 2.4) bilden die drei miteinander verbundene Komponenten epistemischer Zweifel (*epistemic doubt*), epistemische Volition (*epistemic volition*) und Auflösungsstrategien (*resolution strategies*) gemeinsam den Mechanismus der Veränderung epistemologischer Überzeugungen von den aktuell vorherrschenden Überzeugungen (*current beliefs*) hin zu fortgeschritteneren Überzeugungen (*advanced beliefs*).

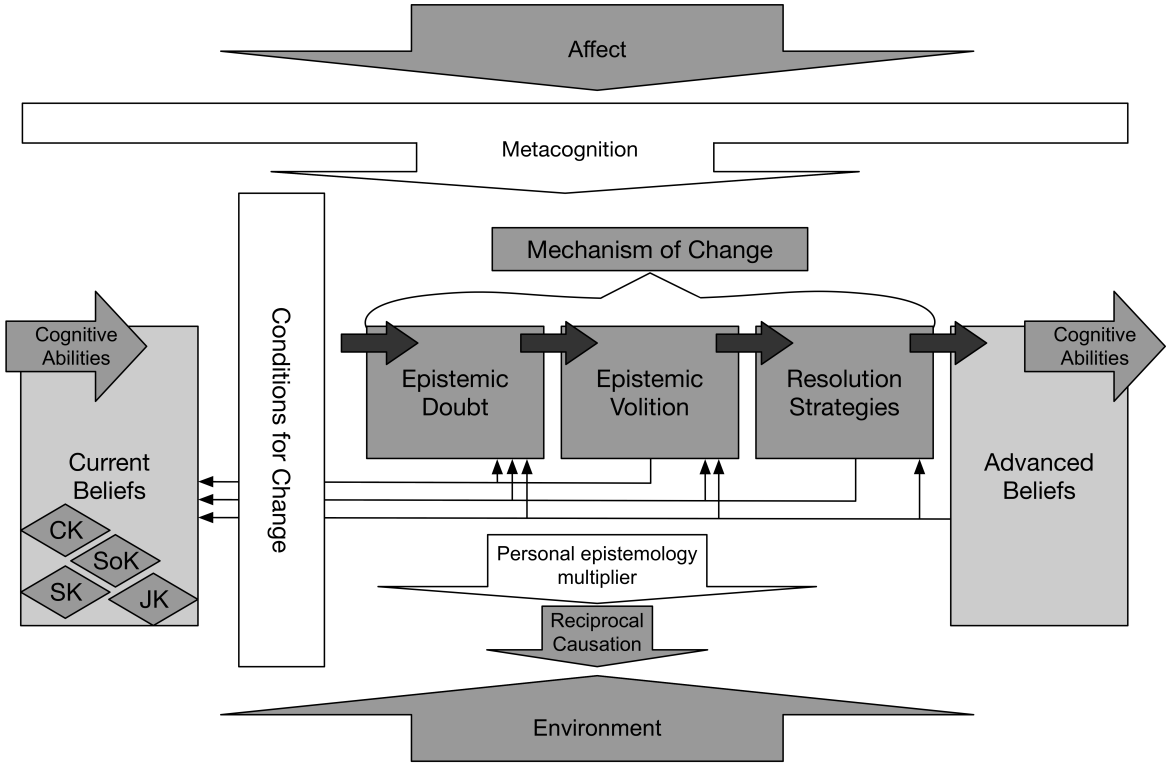


ABBILDUNG 2.4: Mechanismen der Veränderung im *Integrative Model for Personal Epistemology* (IM) von Bendixen und Rule (2004) mit *personal epistemology multiplier* (Rule, 2003) nach Rule und Bendixen (2010, S. 98), in eigener Darstellung

Als Voraussetzungen der Veränderungen (*conditions of change*, siehe Abb. 2.4), dem Veränderungsmechanismus vorgeschaltet, sind die *Dissonanz* und *persönliche Relevanz*, die die Entwicklung anstoßen müssen. Die epistemologische Entwicklung (der Veränderungsmechanismus) wird demnach stimuliert, wenn das Individuum Dissonanz empfindet, das heißt, Erfahrungen macht, die mit Erwartungen nicht übereinstimmen. Widersprüchliche Erfahrungen können eine Person demnach dazu bringen, vorherige Annahmen und Überzeugungen zu überdenken, neu zu interpretieren oder abzulehnen (King & Kitchener, 1994). Übereinstimmend mit den Annahmen zu *Conceptual Change* (Pintrich & Garcia, 1993) muss eine Person zunächst den Eindruck haben, dass ihre aktuellen Überzeugungen nicht mehr zufriedenstellend sind, bevor Veränderungen eingeleitet werden können (Bendixen & Rule,

2004). Damit einher geht die zweite Voraussetzung der Veränderung, die persönliche Relevanz. Widersprüchliche Erfahrungen oder Informationen müssen nicht unbedingt zu einer Veränderung der individuellen Überzeugungen führen. Vielmehr kann es vorkommen, dass das Indiviuum sie ignoriert, wenn die nicht in irgendeiner Form als persönlich bedeutsam empfunden werden (ebd.).

*Epistemic doubt* stellt die erste Komponente des Veränderungsmechanismus dar (siehe Abb. 2.4, S. 58). Dabei handelt es sich um das Anzweifeln der eigenen Überzeugungen (Bendixen & Rule, 2004). Vergleichbar mit dem Modell des *Conceptual Change* von Dole und Sinatra (1998) umfasst das Anzweifeln der eigenen epistemologischen Überzeugungen das Abwägen von Hinweisen, die für andere, konkurrierende Überzeugungen sprechen sowie das Beurteilen der Richtigkeit dieser neuen Überzeugungen (Bendixen & Rule, 2004). Erst wenn die Einschätzung der anderen Überzeugungen dazu führt, dass diese vertrauenswürdig erscheinen, können sich fortgeschrittenere Überzeugungen entwickeln. Die tatsächliche Weiterentwicklung erfordert allerdings nicht nur das Anzweifeln der bisherigen Überzeugungen, sondern auch *epistemic volition* (siehe Abb. 2.4, S. 58). Volition wird dabei nach Corno (1993, S. 16) als ein dynamisches System von psychologischen Kontrollprozessen definiert, das die Konzentration und zielgerichtete Anstrengung vor persönlichen oder umweltbedingten Ablenkungen abschirmt (Bendixen & Rule, 2004, S. 72). *Epistemic volition* „focuses attention on metacognitive awareness and on individuals taking „responsibility“ for their epistemological beliefs “ (Bendixen & Rule, 2004, S. 73). Die Entscheidung zur Veränderung der epistemologischen Überzeugungen muss also metakognitiv gesteuert und abgeschirmt werden vor möglichen Ablenkungen, damit diese Entscheidung umgesetzt werden kann. Abhängig von den beiden vorherigen Komponenten ist die dritte und letzte Komponente des Veränderungsmechanismus die *resolution strategies* (siehe Abb. 2.4, S. 58). Reflexion und soziale Interaktion werden hier als zentrale Strategien des Umgangs mit dem Zweifel an den eigenen Überzeugungen und der Auflösung dieses Zweifels durch eine Weiterentwicklung der epistemologischen Überzeugungen beschrieben (Bendixen & Rule, 2004). Die drei Komponenten sind voneinander abhängig und werden der Reihe nach durchlaufen. Das führt zu einer Weiterentwicklung der epistemologischen Überzeugungen zu den *advanced beliefs*. *Advanced* steht hier jedoch nicht für den Endpunkt der epistemologischen Entwicklung, sondern bedeutet in diesem Zusammenhang nur, dass die Überzeugungen sich (in diesem einen Durchlauf des Veränderungsmechanismus) im Vergleich zu den *current beliefs* weiterentwickelt haben. Nun stellen die *advanced beliefs* die neuen *current beliefs* dar und können wiederum weiterentwickelt werden. Allerdings ist es in jeder Phase des Veränderungsprozesses möglich, dass das Individuum zu den vorher vorherrschenden Überzeugungen zurückkehrt und eine dauerhafte Weiterentwicklung zu fortgeschrittenen Überzeugungen nicht gelingt. Dies gilt auch, nachdem die fortgeschritteneren Überzeugungen eigentlich bereits erreicht wurden. Selbst hier ist eine Rekursion zu den vorherigen Überzeugungen noch möglich (siehe Abb. 2.4, S. 58, Bendixen & Rule, 2004).

Die Metakognition stellt in diesem Modell als exekutiver Kontrollprozess des gesamten Veränderungsmechanismus und der epistemologischen Überzeugungen einen zentralen Bestandteil dar (siehe Abb. 2.4, S. 58). Der Grad der metakognitiven Auseinandersetzung, den

eine Person während ihres epistemologischen Entwicklungsprozesses ausübt, bestimmt, ob die Entwicklung gelingt und von Dauer ist. Je mehr sich ein Individuum metakognitiv mit den fortgeschrittenen epistemologischen Überzeugungen, den Argumenten und Hinweisen, die für diese neuen Überzeugungen sprechen und dem eigenen Zweifel an den aktuellen Überzeugungen auseinandersetzt, desto langlebiger wird dessen epistemologische Entwicklung sein (Bendixen & Rule, 2004).

Bendixen und Rule (2004) postulieren eine dynamische Interaktion der individuellen kognitiven Fähigkeiten (*cognitive ability*) mit den Anforderungen des Umfeldes (z.B. durch Lehrende, Peers), die zur epistemologischen Entwicklung führt. Im Bezug auf das Umfeld wird die zentrale Rolle der Peers betont, die diese bei der individuellen epistemologischen Entwicklung einnehmen. Die Interaktion mit Peers sorgt dafür, dass Bedingungen der Veränderung (*conditions for change*), wie die Dissonanz, eintreten (Bendixen, 2002). Den Peers wird beim komplexen Problemlösen nachgeeifert, sie sind Schlüssel zur Auflösung epistemischer Zweifel (Bendixen, 2002) und nach Schommer-Aikins (2004) Sprungbrett der Veränderung epistemologischer Überzeugungen (vgl. Bendixen & Rule, 2004).

Im Sinne einer Feedbackschleife (*reciprocal causation*, siehe Abb. 2.4, S. 58) beeinflussen sich die individuellen epistemologischen Überzeugungen, die epistemologischen Überzeugungen von Personen im Umfeld und das Umfeld gegenseitig. Bendixen und Rule (2004, S. 76) beschreiben diesen Prozess folgendermaßen: Wenn ein externer Faktor dazu führt, dass die epistemologischen Überzeugungen einer Person sich weiterentwickeln, stellt das eine positive Veränderunge im Umfeld vieler anderer Personen dar, deren epistemologische Überzeugungen sich deshalb unter Umständen ebenfalls weiterentwickeln und so die Qualität des Umfeldes der ersten Person anheben (vgl. *personal epistemology multiplier*, Rule, 2003).

Individuelle Emotionen (*Affect*, siehe Abb. 2.4, S. 58) können die individuelle epistemologische Entwicklung nach Bendixen und Rule (2004) ebenfalls beeinflussen. Eine Person kann sich zum Beispiel aufgrund von starken Gefühlen bezüglich der aktuellen epistemologischen Überzeugung dazu entscheiden, einen empfundenen epistemischen Zweifel zu ignorieren, oder sich beispielsweise dazu entscheiden, Zweifel zu ignorieren, weil sie die Auseinandersetzung scheut (Bendixen, 2002). Epistemologische Veränderung geht mit negativen Emotionen, wie Trauer, Einsamkeit, Sorge, Angst und Verwirrung einher, die im Zusammenhang mit dem *epistemic doubt* auftreten (Bendixen, 2002; Chandler, 1987).

Übereinstimmend mit dem Modell von Bendixen und Rule (2004) beziehen sich die aktuellen Ansätze dazu, wie eine Veränderung der epistemologischen Überzeugungen von Personen gefördert werden kann, entweder auf das Hervorrufen von kognitivem Konflikt und Zweifel (im Sinne des *Conceptual Change*) oder auf die explizite Reflexion der wissensbezogenen Überzeugungen (vgl. Brownlee, Ferguson & Ryan, 2017; Brownlee et al., 2016). Eine Auswahl dieser Interventionen zur Veränderung der epistemologischen Überzeugungen wird im folgenden Abschnitt dargestellt.

**Interventionen zur Veränderung epistemologischer Überzeugungen**

Eine Reihe von sowohl kurz- und langfristigen Interventionen konnten zeigen, dass mittels kognitivem Konflikt und Zweifel eine Veränderung der epistemologischen Überzeugungen von Lehramtsstudierenden erreicht werden kann (für ein Review siehe Brownlee et al., 2016, vgl. Brownlee, Ferguson & Ryan, 2017). In einigen längsschnittlichen Studien wurden die Veränderungen in den wissensbezogenen Überzeugungen von Lehramtsstudierenden während ihres Verlaufs durch die Lehrerbildung erforscht. Hier zeigt sich, dass kognitive Konflikte, die über konstruktivistische Lernansätze (z.B. Rodríguez & Cano, 2007) oder die Konfrontation mit konfligierenden Theorien (z.B. Sosu & Gray, 2012; Walker, Brownlee, Whiteford, Exley & Woods, 2012) hervorgerufen werden, zu Veränderungen der epistemologischen Überzeugungen führen können (Brownlee, Ferguson & Ryan, 2017).

Sosu und Gray (2012) gelang es, Veränderungen in den Überzeugungen zur Fähigkeit des Lernens, zum Lernprozess sowie zur Quelle und Sicherheit des Wissens bei schottischen Lehramtsstudierenden über einen Zeitraum von vier Jahren ($N$ = 112 im ersten Jahr und $N$ = 71 im vierten Jahr) zu zeigen, indem die Studierenden im Rahmen eines Kurses mit verschiedenen Perspektiven zum Lehren und Lernen konfrontiert wurden (Induzierung einer kognitiven Dissonanz). Übereinstimmend identifizierten Walker et al. (2012) im Rahmen von Interviews mit 20 Lehramtstudierenden, deren epistemologische Überzeugungen sich im Laufe eines vierjährigen Bachelors of Education weiterentwickelt hatten, neben akademischen Erfahrungen, Reifeprozessen und Erfahrungen im Feld auch herausfordernde Erfahrungen und kognitives Disäquilibrium als Veränderungsmechanismen.

Auch sehr kurze Interventionen, die kognitive Konflikte über sogenannte *refutational texts* hervorriefen, konnten bereits Veränderungen in den epistemologischen Überzeugungen von Studierenden induzieren (z.B. Brownlee et al., 2016; Gill, Ashton & Algina, 2004; Kienhues, Bromme & Stahl, 2008), auch wenn hier die Langlebigkeit der Veränderungen fraglich ist. *Refutational texts* dienen dazu, *Conceptual Change* anzustoßen, indem sie unter Rückgriff auf wissenschaftliche Evidenz gegensätzliche Perspektiven zu bestimmten Sachverhalten darstellen und so bei Lernenden Unzufriedenheit mit und Zweifel an den eigenen epistemologischen Überzeugungen auslösen (Guzzetti, Snyder, Glass & Gamas, 1993).

Kienhues et al. (2008) wählten für ihre Studie zunächst auf Basis der Ergebnisse eines Pretests insgesamt 58 deutsche Studierende (89% weiblich) aus, von denen die Häfte über naivere und die andere Häfte über sophistiziertere epistemologische Überzeugungen verfügte. Diese Studierenden wurden zufällig auf zwei Gruppen aufgeteilt. Die epistemologischen Überzeugungen der Studierenden in der Experimentalgruppe wurden über einen Text zum Thema DNA-Fingerabdruck, der kontroverse Sichtweisen auf das Thema darstellte, (*refutational epistemological instruction*) herausgefordert, während die Kontrollgruppe einen rein narrativen, informierenden Text erhielt, der keine kontroversen Darstellungen enthielt. Durch den Vergleich der Ergebnisse des Pretest mit einem Posttest wurde der Treatmenteffekt unter Einsatz des CAEB (Stahl & Bromme, 2007) und DEBQ (deutsche Version von Richter, 2004) ermittelt. Die Studierenden mit ursprünglich naiveren epistemologischen

Überzeugungen in der Experimentalgruppe (mit der die bestehenden Überzeugungen herausfordernden Instruktion) veränderten ihre wissensbezogenen Überzeugungen in Richtung größerer Sophistiziertheit. Bei allen anderen Studierenden, sowohl denen mit sophistizierten Überzeugungen der Experimentalgruppe als auch denen mit naiven und sophistizierten Überzeugungen der Kontrollgruppe (informierende Instruktion), zeigte sich eine Veränderung in Richtung naiveren epistemologischen Überzeugungen. Ob die durch das Treatment hervorgerufenen Veränderungen in den bereichspezifischen epistemologischen Überzeugungen von Dauer sind, wurde nicht überprüft und ist äußerst fraglich. Auch wäre eine Veränderung in Richtung naiverer wissensbezogener Überzeugungen, wie sie für die ursprünglich sophistizierteren Studierenden der Experimentalgruppe eintrat, wenig wünschenswert.

Neben dem Ansatz, die Überzeugungsveränderungen durch kognitiven Konflikt zu induzieren, hat sich auch gezeigt, dass eine explizite Reflexion der individuellen wissensbezogenen Überzeugungen zu einer Veränderung ebendieser führen kann (Brownlee et al., 2016). Besonders die metakognitive Reflexion über die eigenen Überzeugungen zum Wissen und Wissensprozess scheint eine Veränderung unterstützen zu können (z.B. Brownlee, Ferguson & Ryan, 2017; Brownlee et al., 2011, 2016; Muis, 2007). Dies konnte beispielsweise für Lehramtsstudierende (z.B. Deniz, 2011; Güven, Sülün & Cam, 2011; Valanides & Angeli, 2005) und für praktizierende Lehrkräfte (z.B. Charalambous, Panaoura & Phillippou, 2009; Tsai, 2006) bestätigt werden. Deniz (2011) beispielsweise ließ 161 Lehramtsstudierende in einem vierwöchigen Kurs schriftlich und im Rahmen von Diskussionen explizit über ihre epistemologischen Überzeugungen reflektieren. Eine signifikante und bedeutende Veränderung der Überzeugungen zeigte sich im Bezug auf alle vier Überzeugungsdimensionen des DEBQ nach Hofer (2000). In der Studie von Valanides und Angeli (2005) wurden 126 US-amerikanische Lehramtsstudierende randomisiert drei Gruppen zugeordnet. Die Interventionen in den drei Gruppen dauerten jeweils 65 Minuten und dienten der Förderung epistemologischer Überzeugungen und der Förderung von Reflexion über das kritische Denken. In der ersten Gruppe wurde das kritische Denken unabhängig vom Fach gelehrt, während die Fertigkeiten zum kritischen Denken in der zweiten Gruppe explizit gelehrt wurden und zusätzlich über das kritische Denken reflektiert wurde. In der dritten Gruppe wurden die Fertigkeiten zum kritischen Denken ebenfalls explizit gelehrt, allerdings wurde die sokratische Methode eingesetzt, um die eigenen Denkfertigkeiten zu evaluieren. Während sich für alle drei Gruppen Veränderungen in den epistemologischen Überzeugungen zeigen ließen, erzielte die zweite Gruppe die größte Veränderung. Weitere Interventionen zur Veränderung der epistemologischen Überzeugungen (mittels Reflexionsprozessen oder *Conceptual Change*) von Studierenden wurden u.a. auch von Marra, Palmer und Litzinger (2000), Muis und Duffy (2013), Kienhues, Stadtler und Bromme (2011), Ferguson et al. (2012) und Gill et al. (2004) umgesetzt. Auch für Schülerinnen und Schüler gibt es zahlreiche Beispiele für Interventionen (z.B. Ferguson & Bråten, 2013; Porsch, 2011). Einen guten Überblick über die Forschung zur Veränderung epistemologischer Überzeugungen mit speziellem Fokus auf angehende und praktizierende Lehrkräfte bieten Brownlee et al. (2016), Brownlee, Ferguson und Ryan (2017), Brownlee, Johansson, Walker und Scholes (2017) sowie Parkinson und

Maggioni (2017).

Es lässt sich festhalten, dass es möglich ist, eine Veränderung der epistemologischen Überzeugungen von Individuen anzustoßen. Dies gelingt, wenn die Lernenden Zweifel und kognitive Konflikte bezüglich der eigenen wissensbezogenen Überzeugungen erfahren oder sich in expliziter, evidenzbasierter Reflexion mit diesen auseinandersetzen (Brownlee et al., 2016). Intensivere, längere Interventionen scheinen dabei zu mehr und langlebigerer Veränderung zu führen als Kurzinterventionen (Elby et al., 2016, S. 117). Für die Entwicklung von sophistizierten epistemologischen Überzeugungen von Lehramtsstudierenden sowie die Überführung dieser Überzeugungen in die Unterrichtspraxis ist es jedoch von Vorteil, wenn die explizite Reflexion eingebettet in die praktische Anwendung stattfindet (vgl. Adibelli-Sahin & Bailey, 2017; Fives & Buehl, 2017; Gholami, 2017; Schraw et al., 2017). Das gibt den Lehramtsstudierenden die Möglichkeit, die eigenen Überzeugungen und Lehrhandlungen zu reflektieren und die konkreten Lehrhandlungen an entstehende sophistiziertere Überzeugungen anzupassen (vgl. Brownlee et al., 2016; Schraw et al., 2017). Es würde sich demnach in der Lehrerbildung anbieten, eine Förderung der epistemologischen Überzeugungen der Lehramtsstudierenden mit Praktika an Schulen zu verbinden.

## 2.4 Explizite Mehrdimensionalität

Die Annahme mehrerer Überzeugungsdimensionen geht auf Schommer (1990) zurück. Hofer und Pintrich (1997) konnten nicht nur zeigen, dass auch die epistemologischen Entwicklungsmodelle spezifische Dimensionen epistemologischer Überzeugungen annehmen, sondern dass über die Modelle hinweg große Übereinstimmungen bezüglich der angenommenen Modelle vorliegen und sich daraus zentrale Kerndimensionen identifizieren ließen (siehe Kap. 2.2.5, Abb. 2.2, S. 44).

In der folgenden Abbildung 2.5 werden die in den verschiedenen theoretischen Konzeptionen angenommenen Dimensionen gegenübergestellt. Das umfasst einerseits die bereits von Hofer und Pintrich (1997) gegenübergestellten Modelle von Perry (1970) bis Schommer (1990), als auch einige neue Konzeptionen. Auch alle hier dargestellten neueren Modelle (siehe Abb. 2.5) nehmen ein System von Überzeugungsdimensionen an. Beim Vergleichen dieser neueren Modelle fällt auf, dass sie alle mit der angenommenen Dimensionalität der Konzeption Hofer und Pintrich (1997) vereinbar sind.

| | Kerndimensionen epistemologischer Überzeugungen | | | | | Periphere Dimensionen epistemologischer Überzeugungen | |
|---|---|---|---|---|---|---|---|
| | Beschaffenheit des Wissens | | | Beschaffenheit des Wissensprozesses | | Beschaffenheit des Lernens | Beschaffenheit der Intelligenz |
| Forschende | Stabilität/ Sicherheit | Struktur/ Einfachheit | Entwicklung | Quelle | Rechtfertigung | Geschwindigkeit des Lernens | Fähigkeit zu lernen |
| Perry (1970) | X | | | X | | | |
| Belenky (1986) | | | | X | | | |
| Baxter Magolda (1992) | X | | | X | X | X | |
| King & Kitchener (1994) | X | X | | X | X | | |
| Kuhn (1991) | X | | | X | X | | |
| Schommer (1990) | X | X | | X | X | X | X |
| Hofer & Pintrich (1997) | X | X | | X | X | | |
| Hofer (2004a) | X | X | | X | X | | |
| Schommer-Aikins (2004) | | X | | X | X | X | X |
| Bendixen & Rule (2004) | X | X | | X | X | | |
| Conley et al. (2004) | X | | X | X | X | | |
| Buehl & Alexander (2006) | X | X | | X | | | |
| Muis et al. (2006) | X | X | | X | X | | |

ABBILDUNG 2.5: Vergleich der angenommenen Dimensionen epistemologischer Überzeugungen zwischen verschiedenen Modellen

In der Benennung der Dimensionen liegen zwischen den Konzeptionen einige kleine Unterschiede vor. Die vielfach als *Einfachheit des Wissens* bezeichnete Dimension, die sich auf die Überzeugungen bezieht, dass Wissen entweder über eine einfache oder aber komplexe Struktur verfügt, wird in den Konzeptionen von Schommer-Aikins (2004) und Buehl und Alexander (2006) als *Struktur des Wissens* bezeichnet. Es handelt sich aber inhaltlich um

die gleiche Dimension. Auch die Dimension, die Hofer als *Sicherheit des Wissens* bezeichnet, heißt bei Schommer (1990) und Buehl und Alexander (2006) zwar *Stabilität des Wissens*, bezieht sich aber auf den gleichen inhaltlichen Aspekt epistemologischer Überzeugungen.

Besonders sei an dieser Stelle auf die Dimensionalität epistemologischer Überzeugungen hingewiesen, die Conley et al. (2004) annehmen. Ihre Forschung basiert auf der Konzeption von Hofer und Pintrich (1997), Hofer (2000) und Elder (2002). Elder (2002) erforschte die epistemologischen Überzeugungen von Fünftklässlern bezüglich den Naturwissenschaften. Sie konnte drei Skalen zu Überzeugungen über *changing nature of knowledge*, *authority as a source of knowledge* und *role of experiments and source of knowledge* identifizieren. Auf Elders (2002) Entwicklungsdimension basiert die entsprechende Dimension bei Conley et al. (2004). Conley et al. (2004) klassifizieren in der Kategorie *Beschaffenheit des Wissens* die beiden Dimensionen *Sicherheit des Wissens* und *Entwicklung des Wissens*, statt wie Hofer (2000) und Hofer und Pintrich (1997) *Sicherheit* und *Einfachheit des Wissens*. Die Dimension *Entwicklung des Wissens* bezieht sich auf die Überzeugung, dass es sich bei der betreffenden Diszplin um einen sich weiterentwickelnden Bereich handelt und dass Ideen und Theorien sich auf der Basis neuer Forschungserkenntnisse verändern können. Personen mit fortgeschritteneren epistemologischen Überzeugungen nehmen an, dass sich die Theorien, Annahmen und Ideen der Disziplin weiterentwickeln und dass Entdeckungen dazu führen, dass die Experten des Bereiches ihre Ansichten ändern. Personen mit weniger fortgeschrittenen epistemologischen Überzeugungen nehmen dagegen an, dass die Theorien, Annahmen und Ideen der Disziplin gleich bleiben, d.h. sich nicht weiterentwickeln (Conley et al., 2004).

Für all diese zuvor postulierten Dimensionen liegen empirische Evidenzen vor. Eine exzellente Übersicht über Studien, in denen im englischsprachigen Raum der SEQ und Varianten desselben eingesetzt wurden, sowie die eingesetzten Auswertungsverfahren und identifizierten Dimensionen, findet sich bei Buehl (2008). Schommer (1990) konnte unter Einsatz des SEQ faktorenanalytisch nur vier der fünf Dimensionen (alle außer *Quelle des Wissens*) anhand einer Stichprobe identifizieren, die 116 College Studierende (58 weiblich, 58 männlich) und 76 Universitätsstudierende (62 weiblich, 4 männlich) umfasste. Diese vierfaktorielle Struktur konnte später mittels exploratorischer und konfirmatorischer Faktorenanalyse anhand neuer Stichproben von 400 Studierenden (*undergraduate*) (Schommer et al., 1992) und 1200 Highschool Schülerinnen und Schülern (Schommer & Dunnell, 2011) repliziert werden (vgl. Schommer, 1993a). Anderen Forschergruppen gelang es auf Basis Schommers (1990) Konzeption, unter Einsatz des SEQ, die vier Dimensionen (simple knowledge, certain knowledge, fixed ability und quick learning) zu replizieren (Bendixen et al., 1994). Weitere Forschergruppen identifizierten faktorenanalytisch mittels modifizierter, gekürzter SEQ Versionen wiederum die vier Skalen *quick learning*, *certain knowledge*, *innate ability*, *depend on authority* (z.B. Chan & Elliott, 2002; Kardash & Scholes, 1996). Anderen gelang es mittels modifizierter und gekürzter SEQ-Versionen die ursprünglichen fünf Dimensionen empirisch zu finden (z.B. Bendixen, Schraw & Dunkle, 1998; Jehng et al., 1993).

Schraw und Kollegen entwickelten auf Basis des SEQ den *Epistemic Belief Inventory* (EBI, Schraw et al., 1995). Faktorenanalytische Auswertungen zu den Daten zweier Stichproben

von 212 College-Studierenden sowie 124 College- und Universitätsstudierenden führten zu fünf reliablen Faktoren, die Schommers Dimensionen entsprechen. Diese konnten in darauffolgenden Studien repliziert werden (Bendixen et al., 1998; Schraw et al., 2002). Schommer-Aikins und Kollegen fanden später bei Schülerinnen und Schülern der siebten bis achten Jahrgangsstufe die Dimensionen *speed of learning*, *ability to learn* und *stability of knowledge* (Schommer-Aikins, Mau, Brookhart & Hutter, 2000) sowie *quick/fixed learning*, *studying aimlessly*, *omniscient authority*, *certain knowledge* (Schommer-Aikins, Duell & Hutter, 2005). Es lässt sich erkennen, dass die auf Basis von Schommers Konzeption gefundenen Dimensionen variieren, jedoch die Dimensionen *Sicherheit/Stabilität des Wissens* und *Struktur des Wissens* durchweg vorkommen. Wie sich in der Übersicht auch zeigt (siehe Abb. 2.5, S. 64), kommen die Dimensionen zur Beschaffenheit der Intelligenz und des Lernens nur in einzelnen Konzeptionen und meist bei Schommer vor. Es handelt sich um periphere Dimensionen, die nicht zum zentralen Konstrukt der epistemologischen Überzeugungen gehören, da sie sich nicht auf das Wissen beziehen (Hofer & Pintrich, 1997). Die Überzeugung eines Lerners dazu, ob Lernen schnell oder gar nicht bzw. langsam vonstatten geht (*Quick Learning*), ist nach Hofer und Pintrich (1997) u.a. Ausdruck der wahrgenommenen Aufgabenschwierigkeit als eine Überzeugung zur Natur des Wissens. Auch die individuelle Überzeugung darüber, ob die Fähigkeit etwas zu Lernen angeboren ist oder aber gelernt werden kann (*Fixed Ability*), betrifft vielmehr „the nature of intelligence as a personal, psychological trait of an individual“ (Hofer & Pintrich, 1997, S. 109).

Hofer (2000) identifizierte auf Basis ihrer Konzeption (Hofer & Pintrich, 1997) unter Einsatz des DEBQ bei 326 Studierenden (*undergraduate*) die Dimensionen *certain/simple knowledge*, *justification for knowing*, *source of knowledge* und *attainability of truth*. Dies entspricht in zwei Aspekten nicht ihrer angenommenen Struktur. Die angenommenen Dimensionen *Sicherheit* und *Einfachheit* bildeten hier eine gemeinsame Skala, während mit *attainability of truth* eine neue Skala identifiziert werden konnte. Conley et al. (2004) konnten die von ihnen auf der Konzeption von Hofer und Pintrich (1997) und Elder (2002) angenommenen vier Skalen (*Sicherheit, Entwicklung, Quelle, Rechtfertigung*) anhand den Daten einer Stichprobe von 187 Fünftklässlern mittels konfirmatorischer Faktorenanalyse bestätigen (siehe Abb. 2.5). In einer deutschsprachigen Übersetzung des Instruments von Conley et al. (2004) gelang es, die Faktorstruktur bei einer Stichprobe von 167 deutschen Schülerinnen und Schülern der neunten Klasse (Urhahne & Hopf, 2004) sowie 120 Lehramtsstudierenden (Biologie) (Urhahne & Harms, 2006) zu replizieren. Der Einsatz qualitativer Erhebungsmethoden, wie Beobachtung und Interviews (Hofer, 2004b) und Laut-Denken-Protokollen (Hofer, 2004a), deutet darauf hin, dass alle vier Kerndimensionen Hofers (Hofer & Pintrich, 1997) (*Quelle, Rechtfertigung, Sicherheit, Einfachheit*) im epistemischen Denken von Studierenden repräsentiert sind. Bråten und Strømsø (2010) entwickelten auf Basis von Hofers vierdimensionaler Konzeption (Hofer & Pintrich, 1997) den *Topic-Specific Epistemic Belief Questionnaire* (TSEBQ) und konnten die angenommene Struktur replizieren. King und Kitchener (1994) und Kuhn (1991) entwickelten ihre Konzeptionen auf Basis der Antworten, die Studierende zu komplexen Problemen gaben (siehe Kap. 2.2.2, ab S. 31 und Kap. 2.2.3, S. 35). Über ihr qualitatives Vorgehen konnten sie verschiedene Entwicklungsstufen epistemologischen Denkens

ableiten. Die Stufen unterscheiden sich jeweils auch im Hinblick auf die Überzeugungen zur *Stabilität des Wissens*, *Struktur des Wissens* (bei King & Kitchener, 1994) und *Quelle des Wissens* und *Begründung des Wissens* über die eine Person verfügt. Auch die Studien, auf deren Basis die Entwicklungsstufenmodelle entstanden, liefern also empirische Evidenz für die Existenz verschiedener Überzeugungsdimensionen.

Wie die oben dargestellte Forschungslage zeigt, gilt die mehrdimensionale Struktur epistemologischer Überzeugungen weithin als angenommen. Es zeigt sich, auch anhand des Aufbaus der neueren Modelle, dass die von Hofer und Pintrich (1997) identifizierten vier Kerndimensionen auch heute als die zentralen Aspekte des Konstrukts epistemologischer Überzeugungen gelten können.

## 2.5 Domänenspezifität und Kontextabhängigkeit

Die Entwicklungsstufenmodelle der ersten Forschungsphase nahmen implizit an, dass individuelle epistemologische Überzeugungen über Disziplinen und Urteilsbereiche hinweg gleich sind (vgl. Muis et al., 2006). Die zwischen Expertinnen und Experten verschiedener Disziplinen gefundenen Unterschiede (Royce, 1978; Royce & Mos, 1980) ließen jedoch Zweifel an intraindividuellen, disziplinübergreifenden epistemologischen Überzeugungen aufkommen (z.B. Schommer, 1993a). Zumal die Klärung dieser Frage einerseits zum Verständnis des Konstrukts beiträgt und auch für die Entwicklung geeigneter Erhebungsinstrumente, wie Fragebögen, von zentraler Bedeutung ist (vgl. Hofer, 2000; Schraw, 2001).

Der Forschungsdiskurs zur Domänenspezifität epistemologischer Überzeugungen begann in der zweiten Forschungswelle, wurde in der dritten Forschungswelle intensiv beforscht und konnte schließlich 2006 weitgehend abgeschlossen werden. Im Folgenden werden die verschiedenen Diskussionsstränge zu dieser Thematik kurz dargestellt. Im Anschluss wird anhand des theoretischen Rahmenmodells *Theory of Integrated Domains in Epistemology (TIDE) Framework* aufgezeigt, welche aktuellen Ansichten zu domänenübergreifenden bzw. domänenspezifischen epistemologischen Überzeugungen bestehen.

Wie in Kapitel 2.1.3 bereits erwähnt, wird der Begriff *Domäne* im Forschungsfeld mit unterschiedlicher Bedeutung verwendet (Hofer, 2006b). Ganz generell ist darunter die Wissensdomäne zu verstehen, auf die sich die individuellen epistemologischen Überzeugungen beziehen. Nach Hofer (2006b) verwenden pädagogische Psychologen diesen Begriff gleichbedeutend mit der akademischen Disziplin (z.B. Muis et al., 2006), wie zum Beispiel Überzeugungen zu Wissen in Geschichte oder Psychologie, während Entwicklungspsychologen darunter eher Urteilsbereiche wie Ästethik und Werte verstehen (z.B. Kuhn, 1991). Gleichzeitig wurde der Begriff Domäne auch schon auf verschiedene Lebensbereiche bezogen, wie Gesundheit, Arbeit und Familie (Pintrich, 2002). Für eine Darstellung der Geschichte des Begriffs *Domain* in Bezug auf Wissen im Bildungsbereich sei an dieser Stelle auf das Review von Stevens, Wineburg, Herrenkohl und Bell (2005) verwiesen.

In dieser Arbeit wird der begrifflichen Verwendung der pädagogischen Psychologie gefolgt, da die akademische Fachdisziplin hier von entscheidender Bedeutung ist. Dem Vorschlag von Hofer (2006b) entsprechend, wird in dieser Arbeit statt des Begriffs *Domäne*, der nicht eindeutig ist, der Begriff *Disziplin* verwendet. Damit ist im Sinne Muis et al. (2006) die akademische Fachdisziplin gemeint, wie beispielsweise die Erziehungswissenschaft.

In Bezug auf die Disziplinspezifität lassen sich zwei zentrale Diskussionsstränge unterscheiden. In den ersten Bereich fällt die Forschung, in der mittels *between-subject* Forschungsdesign untersucht wird, ob sich die Studierenden oder Dozierende verschiedener Disziplinen in ihren epistemologischen Überzeugungen unterscheiden. Auf der anderen Seite besteht Interesse an intraindividuellen Unterschieden in den epistemologischen Überzeugungen (in den Ausprägungen der Überzeugungsdimensionen), die zu verschiedenen Disziplinen vorliegen (Hofer, 2006b). Diese werden mittels *within-subject* Designs untersucht.

Es kann davon ausgegangen werden, dass bestimmte Disziplinen sich in Bezug auf ihre

Struktur und Inhalte (Alexander, 1992; Frederiksen, 1984; Spiro & Jehng, 1990) sowie die vorherrschenden Kriterien zur Begründung und Rechtfertigung des Wissens (Royce, 1959, 1978; Royce & Mos, 1980) unterscheiden (vgl. Buehl & Alexander, 2001; Muis et al., 2006).

Schon Royce (1959) nahmen die Existenz verschiedener Wissensmodi an. Er ging davon aus, dass in den verschiedenen Disziplinen (z.B. Kunst vs. Naturwissenschaft) unterschiedliche Kriterien zur Begründung und Rechtfertigung des Wissens vorherrschen. Im Laufe der Sozialisation in den verschiedenen Disziplinen, die mit der akademischen Bildung einhergeht, nähern sich Individuen in ihrem epistemologischen Denken dem in der Disziplin vorherrschenden an. Je nach Disziplin dominiert einer von drei grundlegenden kognitiven Prozessen den Wissenserwerb (*epistemic hierarchy*) (Royce, 1978). Bei diesen sogenannten *Ways of Knowing* handelt es sich um den Rationalismus (*knowing via thought*), Empirismus (*knowing via the senses*) und Metaphorismus (*knowing via intuition*). Mittels des standardisierten Inventars *Psycho-Epistemological Profile* (PEP, Royce & Mos, 1980) lassen sich die kognitiven Präferenzen und das epistemologische Überzeugungssystem von Individuen empirisch erfassen. Je nach individueller Ausprägung bezüglich der drei *Ways of Knowing* ergibt sich eine unterschiedliche Gewichtung, die das psycho-epistemologische Profil einer Person darstellt (Royce, 1978; Royce & Mos, 1980). Bei überwiegendem Rationalismus eignet sich eine Person Wissen hauptsächlich durch logisches und analytisches Denken an. Wird Wissen hauptsächlich durch die Analyse von Beobachtungen und Daten und deren Bedeutung gewonnen, dominiert der Empirismus. Im Metaphorismus wird Wissen als subjektiv betrachtet und der Wissenserwerb durch die Intuition bestimmt (Royce, 1978).

Zur Prüfung der These, dass in verschiedenen Disziplinen unterschiedliche Kriterien zur Begründung und Rechtfertigung des Wissens vorherrschen, wählten Royce und Mos (1980) Universitätsprofessorinnen und Professoren als Experten und Vertreter der Disziplinen Litertur und Kunst, Mathematik, theoretische Physik und Naturwissenschaften aus. Über den Einsatz des PEP konnten sie zeigen, dass sich die in den jeweiligen Disziplinen vorherrschenden epistemologischen Überzeugungen unterscheiden. Im literarisch-künstlerischen Bereich war der Metamorphismus die dominante epistemologische Denkweise unter den Professorinnen und -professoren, während in Mathematik und theoretischer Physik der Rationalismus und in den Naturwissenschaften der Empirismus vorherrschte (Muis et al., 2006). Donald (1990) verglich die Sichtweisen von Universitätsprofessoren verschiedener Fachbereiche bezüglich der geeigneten Kriterien zur Rechtfertigung und Begründung des Wissens. Sie fanden heraus, dass Professoren der Natur- und Sozialwissenschaften hauptsächlich empirische Evidenzen, logische Konsistenz und Reliabilität als geeignete Kriterien betrachten, Wissen zu begründen und als richtig zu erachten. In den Geisteswissenschaften wurde dagegen bei der Wissensbegründung eher auf die Autorität anderer vertraut. Nach Biglan (1973a, 1973b) lassen sich akademische Disziplinen danach unterscheiden, ob sie erstens über ein einheitliches zugrundeliegendes Paradigma verfügen (paradigmatischer

= *hart*, weniger paradigmatisch = *weich*), sich zweitens mit der praktischen Anwendung befassen (theoretische Auseinandersetzung = *rein*, theoretische Auseinandersetzung mit praktischer Anwendung = *angewandt*) und sich drittens mit lebenden Systemen beschäftigen. Paradigmatischere, *harte* Fächer sind demnach beispielsweise Physik oder Mathematik, während die Erziehungswissenschaft und die Geisteswissenschaften als weniger paradigmatisch und demnach *weich* klassifiziert werden, da ihre Inhalte und angewandten Methodologien idiosynkratischer seien. Disziplinen, die auf die theoretische Entwicklung fokussiert sind und sich nicht mit der Anwendung beschäftigen, wie theoretische Physik, werden nach dieser Klassifikation als *rein* bezeichnet, während die anderen Disziplinen, wie Ökonomie, als *angewandt* bezeichnet werden (vgl. Muis et al., 2006, S. 10 f.). Die Disziplin Biologie befasst sich beispielweise mit lebenden Systemen, während Physik sich mit nicht-lebenden Systemen beschäftigt. Diese Klassifikationen sind allerdings nicht immer eindeutig. Mathematik wird nach diesem Klassifikationsschema beispielsweise als *rein* und *hart* kategorisiert (vgl. Muis et al., 2006, S. 10 f.), allerdings würde der Teilbereich der Angewandten Mathematik nicht unter *rein* sondern *angewandt* eingeordnet werden. Die Erziehungswissenschaft als Disziplin lässt sich nach Royce (1978) den oben dargestellten Klassifikationssystemen als schwach strukturiert, weich und angewandt beschreiben. Die vorherrschende epistemologische Denkweise ist unter Fachexpertinnen und Fachexperten der Empirismus.

Nach Kitchener (1983) können akademische Disziplinen auf Basis der in ihnen am häufigsten vorkommenden Problemstellungen in gut und schlecht strukturiert unterteilt werden. Auch Frederiksen (1984) und Stewart (1987) wählten diese Unterteilung. Allerdings bezieht sie sich bei ihnen auf die verfügbaren Lösungsansätze einer Disziplin. Mathematik und Physik werden demnach den gut strukturierten Disziplinen zugeteilt, weil sie über einheitlich angewandte Lösungsstrategien (abgeleitet aus algorithmischen Prozeduren) verfügen. Disziplinen wie Geschichte jedoch sind schlecht strukturiert, da Problemstellungen hier eher über heuristische Prozeduren gelöst werden (Spiro, Coulson, Feltovich & Anderson, 1994; Wineburg, 1996). Die Problemlösungen in schlecht strukturierten Disziplinen erfordern nach Spiro, Feltovich und Coulson (1996) größere Flexibilität im Denken, da diese Disziplinen multidimensionaler seien. Aufgrund der unterschiedlichen inhärenten Natur (Schwab, 1964) und/oder unterschiedlichen Lehrkonventionen (Stodolsky, 1988) von gut und schlecht strukturierten Disziplinen, kann angenommen werden, dass sich die Verarbeitung und die Arten des Wissens unterscheiden, die Problemlösungen in diesen Disziplinen erfordern (vgl. Buehl & Alexander, 2001, S. 402). Einige Studien (z.B. Jehng et al., 1993; Lonka & Lindblom-Ylänne, 1996; Paulsen, 1998; Schommer, 1993a) fanden Designs Unterschiede in den epistemologischen Überzeugungen zwischen Studierendengruppen verschiedener Disziplinen (Muis et al., 2006).

Jehng et al. (1993) beispielsweise nutzen Biglans Klassifikationssystem, um Unterschiede

in den epistemologischen Überzeugungen zwischen 398 Studierenden (*graduate*, *undergraduate*) *weicher, schlecht strukturierter* Disziplinen (Geisteswissenschaften, Kunst, Sozialwissenschaften) und *harter, gut strukturierter* Disziplinen (Ingenieurswissenschaft, Naturwissenschaften, Business) zu untersuchen. Zur mehrdimensionalen Erfassung der epistemologischen Überzeugungen entwickelten sie auf der Basis von Schommers (1990) Konzeption und den Items des SEQ eine modifizierte Version des SEQ. Die Dimensionen *certainty of knowledge*, *sources of knowing*, *orderly process of learning* (ähnlich zu *simple knowledge*), *innate ability* und *quick learning* konnten faktorenanalytisch (CFA) bestätigt werden. Mittels multivariater Varianzanalyse (MANOVA) fanden sie signifikante Mittelwertsunterschiede zwischen den Studierendengruppen der beiden Disziplinbereiche bezüglich drei der Überzeugungsdimensionen. Studierende der *harten* Disziplinen verfügten über signifikant „naivere" Überzeugungen in Bezug auf *Quelle und Sicherheit des Wissens* sowie *orderly process of knowledge* als Studierende der *weichen* Disziplinen. Studierende *harter* Disziplinen betrachteten Wissen also eher sicher, Autoritäten als Wissensquelle und den Lernprozess als einfach. Signifikante Unterschiede in den epistemologischen Überzeugungen fanden sie auch zwischen den *undergraduate* und den *graduate* Studierenden. Jehng et al. (1993) schlussfolgerten daraus, dass die individuellen epistemologischen Überzeugungen von der Disziplin abhängig sind und sich im Verlauf der akademischen Ausbildung verändern.

Schommer (1993a) fand Unterschiede in den epistemologischen Überzeugungen von 263 Studierenden (College und Universität) der Bereiche Sozialwissenschaften und Technik. Die Studierenden des Bereichs Technik waren eher der weniger fortgeschrittenen Überzeugung, dass man etwas entweder schnell lernt oder aber gar nicht lernen kann. Lonka und Lindblom-Ylänne (1996) fanden Unterschiede in den epistemologischen Überzeugungen von Psychologie- und Medizinstudierenden. Während erstere hauptsächlich über konstruktivistische Vorstellungen zu Lernen und Wissen verfügen, herrscht unter Medizinstudierenden ein Lernverständnis vor, dass sich auf die Rezeption von Fakten beschränkt. Sie betrachten Wissen eher als Sammlung von unverbunden Aussagen. Trautwein et al. (2004) fanden Unterschiede in den epistemologischen Überzeugungen zwischen insgesamt 108 Studierenden *harter* und *weicher* Fachbereiche. Die Studierenden der Sprach- und Kulturwissenschaften sowie die der Rechts-, Wirtschafts- und Sozialwissenschaften betrachteten die Erkenntnisse ihrer Disziplin als wenig überdauernd, während Studierende der Ingenieur- und Naturwissenschaften und Mathematik sie als deutlich stabiler einschätzten.

Mit den Worten Hofers lässt sich der vorher dargestellte Sachverhalt folgendermaßen zusammenfassen:

> „Academic disciplines differ with regard to standards of evidence and justification, the structure of knowledge, and the likelihood of certitude. Art historians, economists, and geologists do not use similar methodological approaches nor substantiate their knowing in similar ways." (Hofer, 2006b, S. 68).

Akademische Disziplinen unterscheiden sich also bezüglich der Standards der Beweisführung und Begründung von Wissen, der Struktur des Wissens und der Wahrscheinlichkeit sicheren Wissens. In den unterschiedlichen Disziplinen werden jeweils andere methodologische Ansätze und Arten der Begründung von Wissensprozessen genutzt (Hofer, 2006b, S. 68).

Neben den oben genannten Studien, die interindividuelle Unterschiede in den epistemologischen Überzeugungen zeigen konnten, gibt es auch viele Studien, die sich den intraindividuellen disziplinspezifischen epistemologischen Überzeugungen widmen. In einer qualitativen Interviewstudie mit 60 Grundschulschülerinnen und -schülern der fünften Jahrgangsstufe konnte Stodolsky, Salk und Glaessner (1991) intraindividuelle Unterschiede in den Überzeugungen bezüglich Mathematikwissen und Wissen in Sozialwissenschaft zeigen. Auch Schoenfeld (1989) und Mason (2003) fanden in ihren quantitativen Studien Hinweise auf die Disziplinspezifität der epistemologischen Überzeugungen bei Schülerinnen und Schülern. Während Schommer und Walker (1995) sowie Schommer-Aikins (2003) unter Einsatz des *Discipline-focused SEQ* nur Evidenz für disziplinübergreifende (allgemeine) epistemologische Überzeugungen bei Studierenden fanden, gibt es doch viele weitere Studien (z.B. Estes et al., 2003; Hofer, 2000; Mori, 1999) mit Studierenden-Stichproben, die die Disziplinspezifität bestätigen (Muis et al., 2006).

Hofer (2000) fand unter Einsatz des DEBQ bei einer Stichprobe von 326 Psychologiestudierenden (53% Frauen) im erste Jahr des Colleges signifikante intraindividuelle Unterschiede in den epistemologischen Überzeugungen zu Psychologie und Naturwissenschaften. Wissen in Naturwissenschaften wurde als signifikant sicherer und unveränderlicher als in Psychologie betrachtet. In Psychologie sahen die Studierenden eigene Erfahrungen eher als geeignete Grundlage der Rechtfertigung von Wissen als in Naturwissenschaften. Autoritäten und Expertise wurden in Naturwissenschaften stärker als Quelle des Wissens gesehen als in Psychologie. Auch gingen Studierenden der Naturwissenschaften eher davon aus, dass Experten hier in der Lage sind, bezüglich eines Sachverhalts die Wahrheit herauszufinden (*attainability of thruth*) als in Psychologie.

Buehl und Kollegen fanden bei Studierenden jeweils Hinweise auf sowohl disziplinspezifische als auch -übergreifende epistemologische Überzeugungen (Buehl & Alexander, 2004, 2005; Buehl et al., 2002). Buehl et al. (2002) befragten in einer Serie von drei Studien Studierende ($N$ = 182/633/523) zu ihren epistemologischen Überzeugungen in den Disziplinen Mathematik und Geschichte. Dazu nutzten sie bis auf die Fachbezeichnung identisch formulierte Items, die sie faktorenanalytisch prüften. In der ersten Studie fanden sie nur je einen Faktor für Mathematik und für Geschichte. In den beiden folgenden Studien konnte eine vierfaktorielle Struktur mit den Faktoren *need for effort* und *integration of information and problem solving* je für Mathematik und Geschichte bestätigt werden.

Olafson und Schraw (2006) setzten in ihrer Studie zur Erfassung der epistemologischen Überzeugungen (*simple knowledge*, *certain knowledge* und *omniscient authority*) von Grundschullehrkräften sowohl Interviews als auch das *Epistemic Belief Inventory* von Schraw et al. (2002) bezüglich der eingesetzten Lehrmethoden ein. Auch sie fanden Hinweise zu disziplinspezifischen und gleichzeitig -übergreifenden wissensbezogenen Überzeugungen.

Wie oben zu sehen ist, existieren einige empirische Studien, die mittels *within-subject* oder aber *between-subject* Designs der Frage nach der Domänenspezifität epistemologischer Überzeugungen nachgehen. Vielfach konnten intraindividuelle und interindividuelle Unterschiede in den epistemologischen Überzeugungen gezeigt sowie aber gleichzeitig auch die Existenz übergreifender Überzeugungen bestätigt werden. Für einen umfassenderen

Einblick in diesen Forschungsdiskurs sei auf Buehl und Alexander (2001), Hofer (2006a) und Hofer (2006b) sowie die Beiträge zur Domänenspezifität im Sammelband von Khine (2008) verwiesen.

Zwei weitere, neuere Forschungsentwicklungen betreffen eine mögliche Themenspezifität (z.B. Bråten et al., 2008; Stahl & Bromme, 2007; Trautwein et al., 2004; Trautwein & Lüdtke, 2007) und Kontextabhängigkeit epistemologischer Überzeugungen (z.B. Bråten & Strømsø, 2010; Haerle & Bendixen, 2008; Hammer, 1994; Hammer & Elby, 2002; Louca et al., 2004; Maggioni, Fox & Alexander, 2010; Mason, Ariasi & Boldrin, 2011; Muis, 2008; Op 't Eynde, Corte & Verschaffel, 2006; Rosenberg, Hammer & Phelan, 2006; Weinstock & Zviling-Beiser, 2009).

In zwei Studien mit je über 600 Studierenden des Fachs Biologie setzten Stahl und Bromme (2007) ihren *Connotative Aspects of Epistemological Beliefs* (CAEB) ein. Sie befragten die Studierenden zu den Themen Pflanzenbestimmung, Genetik und Physik (Studie 1) sowie Pflanzenbestimmung und organische Chemie (Studie 2). Die faktorenanalytische Auswertung resultierte in einer zweifaktoriellen Struktur mit den Faktoren *Beschaffenheit* und *Variabilität*. Stahl und Bromme (2007) fanden signifikante Unterschiede in den epistemologischen Überzeugungen in Abhängigkeit der fachlichen Themen.

Bezüglich der Kontextabhängigkeit wird nach Grossnickle et al. (2017) beispielsweise der Frage nachgegangen, ob sich stabile Unterschiede in den epistemologischen Überzeugungen zwischen den Kulturen verschiedener Länder oder aber unterschiedlicher Bildungskulturen finden lassen (z.B. Bråten & Strømsø, 2010; Haerle & Bendixen, 2008; Hofer, 2007, 2008; Muis, 2008). Es gibt bereits verschiedenste empirische Hinweise darauf, dass der kulturelle Hintergrund sich tatsächlich auf die epistemologischen Überzeugungen auswirkt (z.B. Chai et al., 2006; Chan & Elliott, 2002; Hofer, 2007; Karabenick, 2005; Qian & Alvermann, 1995; Youn, 2000). In der Studie von Youn (2000) wurden explizit koreanische und amerikanische Studierende verglichen. Hier ergaben sich bei Einsatz des gleichen Erhebungsinstruments (JEQ, Jehng et al., 1993) in zwei sprachlichen Versionen unterschiedliche Faktorstrukturen für die amerikanische und die koreanische Version. Die amerikanische Version entsprach in ihrer Faktorstruktur der des Originals von Jehng et al. (1993), während die koreanische Version abwich. Qian und Alvermann (1995) fanden Unterschiede in den epistemologischen Überzeugungen von US-amerikanischen und chinesischen Schülerinnen und Schüler. Chinesische Schülerinnen und Schüler waren eher der Überzeugung, dass Wissen einfach und sicher ist sowie die Fähigkeit zu lernen angeboren, als die amerikanischen. Die amerikanischen Schülerinnen und Schüler waren dagegen eher der Ansicht, dass Lernen entweder schnell geht, oder gar nicht funktioniert. Interessanterweise erreichte der Faktor *certain-simple* knowledge, der sich in verschiedenen US-amerikanischen Studien immer als sehr stabil und robust herausgestellt hatte, in der chinesischen Version des Fragebogens nur eine sehr niedrige Reliabilität von $\alpha$ = .27 (Hofer, 2008). Auch zwischen US-amerikanischen und japanischen Studierenden wurden bereits Unterschiede in den epistemologischen Überzeugungen gefunden (Hofer, 2007), die einerseits darauf hindeuten, dass die kulturelle Sozialisation, beispielsweise über den Wert, der Autorität zugestanden wird,

einen Einfluss auf epistemologische Überzeugungen ausüben könnte. Andererseits deuten diese Ergebnisse auch an, dass möglichweise vom kulturellen Hintergrund abhängig sein könnte, welche Überzeugungsausprägungen gesellschaftlich als sophistizierter betrachtet werden, da die amerikanischen Studierenden über alle vier Dimensionen hinweg über sophistiziertere Überzeugungen zu Wissen verfügten (Hofer, 2007, 2008).

Auch werden inzwischen handlungsbezogene Kontexte erforscht, d.h. solche, in denen Individuen ihre wissenbezogenen Überzeugungen in Handlungen überführen (Mason et al., 2011; Rosenberg et al., 2006; Weinstock & Zviling-Beiser, 2009). Die ersten empirischen Hinweise auf Kontextabhängigkeit epistemologischer Überzeugungen fand Hammer (1994) in der Disziplin Physik. Er konnte zeigen, dass die individuellen Überzeugungen zu Wissen und dem Wissensprozess vom physikalischen Kontext abhängen. Auf Basis dieser und weiterer Befunde konzeptualisiert diese Forschergruppe (Hammer & Elby, 2002; Louca et al., 2004) epistemologische Überzeugungen als kognitive Struktur von epistemologischen Ressourcen (*epistemological resources*), die dazu dienen Wissen zu verstehen. Der individuelle Kontext (konkrete Aufgabe/Problemstellung) sowie die betreffende Disziplin bestimmen, welche epistemologischen Annahmen zur Sicherheit, Struktur und Begründung von Wissen aus den Grundüberzeugungen konstruiert werden (Hammer & Elby, 2002).

Op 't Eynde et al. (2006) fanden damit übereinstimmend bei 365 flämischen Gymnasiasten im Fach Mathematik empirische Hinweise darauf, dass epistemologische Überzeugungen in einen *sozialen Kontext*, *Kontext des Subjekts* und *Kontext der mathematischen Bildung* eingeordnet werden können.

Nach Einschätzung Hofers (2006a) gelang es Muis und Kollegen 2006 den Diskurs um die Domänenspezifität abzuschließen. Auf der Basis ihres Reviews von 19 empirischen Studien (11 *within-subject*, 9 *between-subject*), die sich alle der Erforschung der Domänenspezifität widmeten, entwickelten Muis et al. (2006) das theoretische Rahmenmodell *Theory of Integrated Domains in Epistemology Framework* (TIDE). Das TIDE Framework vereint die oben dargestellten Stränge der Disziplinspezifität in sich und erklärt, unter Berücksichtigung verschiedener Kontexte, wie sich aus den allgemeinen epistemologischen Überzeugungen einer Person die akademischen und die disziplinspezifischen epistemologischen Überzeugungen entwickeln. Darüber hinaus integriert es die Perspektiven epistemologischer Entwicklung mit der explizit mehrdimensionalen Konzeption von Hofer und Pintrich (1997).

## 2.6 Theory of Integrated Domains in Epistemology (TIDE) Framework

Das TIDE Framework wurde von Muis et al. (2006) auf Basis der Ergebnisse von 19 qualitativen und quantitativen empirischen Studien entwickelt, in denen nicht nur viele unterschiedliche Erhebungsinstrumente (z.B. Interviews; SEQ; *Psycho-epistemological profile*, PEP Royce & Mos, 1980; DSBQ sowie Instrumente von Kuhn & Pearsall, 2000; Perry, 1970; M. P. Ryan, 1984) eingesetzt wurden, sondern denen auch unterschiedliche theoretische Konzeptionen epistemologischer Überzeugungen zugrundeliegen. Im Rahmen des TIDE Frameworks, das in Abbildung 2.6 auf Seite 76 zu sehen ist, werden die epistemologischen Überzeugungen in den soziokulturellen, akademischen und instruktionalen Kontext eingebettet. Muis et al. (2006) integrieren die explizit mehrdimensionale Konzeptualisierung von Hofer und Pintrich (1997) und Hofer (2000) mit Modellen über die epistemologische Entwicklung über die Lebensspanne, wie z.B. King und Kitchener (1994). Wie Bendixen und Rule (2004) (*Integrated Model of Personal Epistemology*) gehen sie davon aus, dass sich das epistemologische Denken entwickelt, mehrdimensional ist und Potenzial für wechselseitige Verursachung zwischen den epistemologischen Überzeugungen, Individuen und den verschiedenen Kontexten besteht (Muis et al., 2006). Wie oben bereits erwähnt, werden im TIDE Framework darüber hinaus explizit die Erkenntnisse zu domänenspezifischen und domänenübergreifenden epistemologischen Überzeugungen und zu der Verbindung zwischen beiden berücksichtigt. Die folgende Beschreibung des TIDE Frameworks basiert auf Muis et al. (2006, S. 30 ff.) sowie Muis et al. (2016).

Epistemologische Überzeugungen werden übereinstimmend mit der Konzeption von Hofer und Pintrich (1997) und Hofer (2000) übergeordnet als Überzeugungen zu Wissen und dem Wissensprozess betrachtet, die sich wiederum in die vier Dimensionen Überzeugungen zur *Quelle des Wissens (source of knowledge)*, *Einfachheit des Wissens (simplicity of knowing)*, *Sicherheit des Wissens (certainty of knowledge)*, *Begründung/Rechtfertigung des Wissens (justification for knowing)* (vgl. Muis et al., 2006, 2016). Jede dieser Dimensionen entwickelt sich von einer weniger konstruktivistischen bis hin zu einer konstruktivistischen Perspektive auf Wissen und den Prozess des Wissens. In ersterer wird Wissen als unveränderbar, unverbunden, von Autoritäten vermittelt und als zweifelsfrei wahr betrachtet, ohne dass eine kritische Auseinandersetzung mit den jeweiligen Wissensquellen nötig sei. Personen der konstruktivistischen Perspektive betrachten Wissen als vorläufig, komplex miteinander verbunden, persönlich konstruiert und kritisch zu hinterfragen (Muis et al., 2016, S. 266).

Es wird im TIDE auf Basis der Ergebnisse der empirischen Studien zwischen drei Arten von Überzeugungen zu Wissen und zum Wissensprozess unterschieden, den allgemeinen, den akademischen und den disziplinspezifischen epistemologischen Überzeugungen. In allen drei finden sich die vier verschiedenen Überzeugungsdimensionen. Epistemologische Überzeugungen werden im TIDE Framework (Muis et al., 2006) insofern im soziokulturellen Kontext situiert, als dass sie als komplex und sozial konstruiert betrachtet werden. Das bedeutet, dass Individuen aktiv die Bedeutung zu ihren Erfahrungen konstruieren.

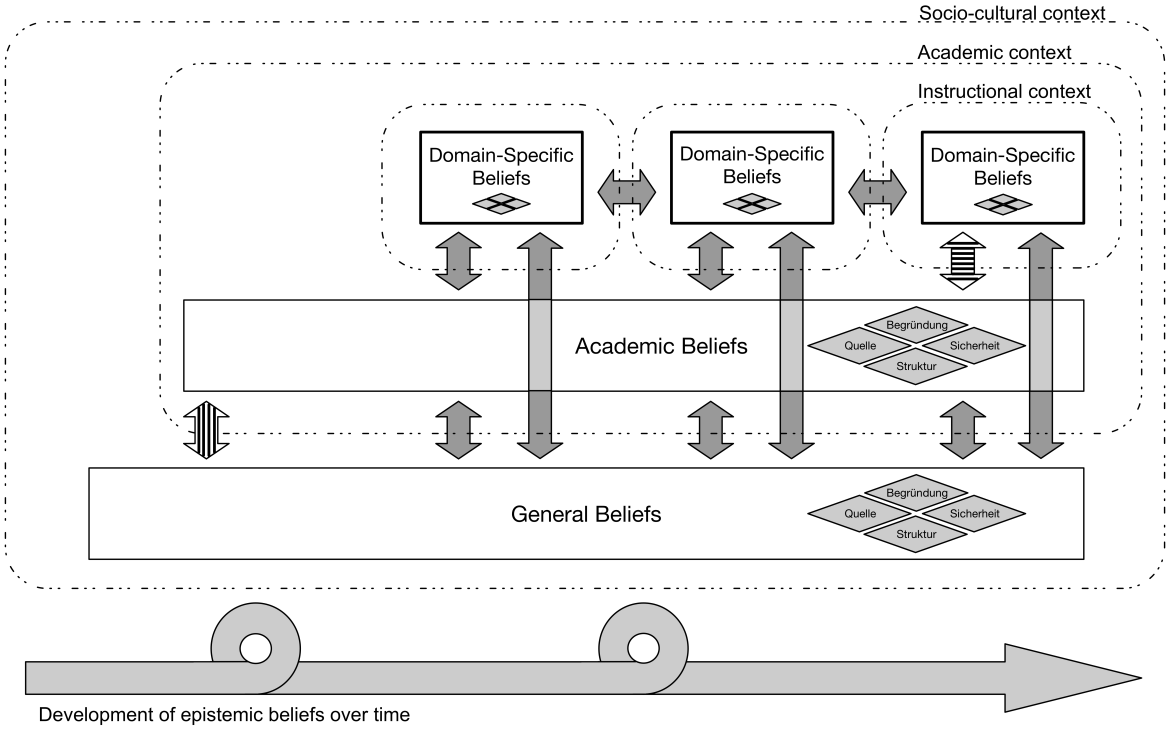


ABBILDUNG 2.6: Überblick über die *Theory of Integrated Domains in Epistemology* (TIDE Framework) nach Muis et al. (2006, S. 30), in eigener Darstellung

Die Entwicklung der allgemeinen epistemologischen Überzeugungen eines Individuums vollzieht sich in der dynamischen Interaktion mit der alltäglichen sozialen und kulturellen Umwelt (Baxter Magolda, 2004; Belenky et al., 1986; Bendixen & Rule, 2004; Hofer & Pintrich, 1997) über die gesamte Lebensspanne (Muis et al., 2006, S. 30 f.). Allgemeine epistemologische Überzeugungen sind nur solche, die in allen nicht akademischen Kontexten entstehen, das heißt im Alltag, zum Beispiel zu Hause, im Arbeitsumfeld und in Interaktion mit Peers (Muis et al., 2006). Die Aneignung der allgemeinen epistemologischen Überzeugungen wird übereinstimmend mit Jehng et al. (1993) als Prozess der Enkulturation verstanden (Muis et al., 2006). Epistemologische Überzeugungen bilden sich im Individuum demnach durch die dynamische Interaktion von kognitiven Kapazitäten mit den Umweltanforderungen (ebd.).

Von diesen allgemeinen epistemologischen Überzeugungen eines Individuums sind die akademischen epistemologischen Überzeugungen zu trennen. Diese sind im akademischen Kontext situiert, das heißt ihre Entwicklung wird durch den Beginn der formalen Bildung des Individuums angestoßen und ganz zentral durch diesen Kontext beeinflusst. Die akademischen epistemologischen Überzeugungen werden vom Individuum im akademischen Umfeld, wie Schule und Hochschule, in dynamischer Interaktion mit diesem Umfeld konstruiert (Muis et al., 2006, S. 31 f.) und beziehen sich als solche ganz zentral auf das Wissen, das speziell in Schule bzw. Universität eine Rolle spielt und dort abgedeckt wird (z.B. mathematisches Wissen, wirtschaftswissenschaftliches Wissen). Aus Abbildung 2.6 geht hervor, dass der akademische Kontext in den soziokulturellen Kontext eingebettet ist (gestrichelte

Linie) und sich die allgemeinen und die akademischen epistemologischen Überzeugungen wechselseitig beeinflussen. In den akademischen Kontext sind wiederum verschiedene instruktionale Kontexte eingebettet. Unter dem instruktionalen Kontext werden die speziellen instruktionalen Bedingungen der betreffenden Bildungseinrichtung und des Bildungssystems verstanden, wie z.B. Notengebung, Schulpolitik und schulische Praktiken sowie Unterricht in verschiedenen Fächern (Schoenfeld, 1989). Mit Eintritt in das Bildungssystem und mit fortschreitendenden Bildungserfahrungen entwickelt das Individuum in den verschiedenen instruktionalen Kontexten, also im Unterricht in verschiedenen Fächern, in verschiedenen Fachbereichen parallel jeweils verschiedene disziplinspezifische epistemologische Überzeugungen (Jehng et al., 1993; Paulsen, 1998; Perry, 1970; Schoenfeld, 1989; Stodolsky et al., 1991), wie spezifische Überzeugungen zum Wissen und dem Wissensprozess für z.B. Mathematik, Musik und Geschichte (dargestellt durch die drei Boxen *domain-specific beliefs*, Abb. 2.6, S. 76).

Auch hier entwickeln sich die disziplinspezifischen epistemologischen Überzeugungen über den Prozess der Enkulturation, d.h. in dynamischer Interaktion mit dem instruktionalen Umfeld. Schülerinnen und Schüler bzw. Studierende lernen demnach das Wissen einer Disziplin näherungsweise so zu betrachten, wie die Personen (z.B. Lehrkräfte, Dozierende) in diesem Kontext (Jehng et al., 1993). Allerdings kommen Lerner in der Sophistiziertheit ihrer Sichtweise auf Wissen sehr lange (bis Abschluss des Studiums) nicht an die von Experten des Fachs heran. Dies wird unter anderem dadurch verursacht, dass das instruktionale Umfeld eine vereinfachte Perspektive auf das disziplinspezifische Wissen vermittelt (Muis et al., 2006, S. 37). Die Mathematikausbildung bedient sich einer weniger sophistizierten Sichtweise auf Mathematikwissen als die Mathematikwissenschaft (Stodolsky et al., 1991).

Auch zwischen den disziplinspezifischen (hier *domain-specific* genannten) Überzeugungen untereinander sowie den disziplinspezifischen und den akademischen epistemologischen Überzeugungen liegen wechselseitige Einflüsse vor. Ebenso wie zwischen den disziplinspezifischen und den allgemeinen epistemologischen Überzeugungen (Muis et al., 2006, S. 31 f.). Mit zunehmender akademischer Erfahrung, d.h. im Verlauf der schulischen Ausbildung und Hochschulbildung, nimmt der Einfluss der allgemeinen epistemologischen Überzeugungen auf die akademischen epistemologischen Überzeugungen ab und die disziplinspezifischen epistemologischen Überzeugungen gewinnen an Bedeutung. Das heißt, zu Beginn der akademischen Ausbildung, im frühen Kindesalter, bestimmen die allgemeinen Überzeugungen noch sehr stark die akademischen Überzeugungen des Individuums (vertikal gestrichelter Pfeil, Abb. 2.6, S. 76), während disziplinspezifische Überzeugungen zuerst noch gar nicht vorliegen. Mit Fortschreiten der akademischen Ausbildung macht das Individuum Bildungserfahrungen in den verschiedensten Bereichen wie Mathematik, Geschichte, etc. und bildet so verschiedene disziplinspezifische epistemologische Überzeugungen aus. Diese üben mit der Zeit einen immer stärkeren Einfluss auf die individuellen akademischen Überzeugungen aus. Im späteren Verlauf der akademischen Ausbildung (vor allem in der Hochschule) spezialisiert sich das Individuum auf eine bestimmte Fachdisziplin, indem es z.B. über die Wahl von Haupt-, Leistungsfächer oder des Studienfachs Schwerpunkte setzt. Daraufhin nimmt der Einfluss der epistemologischen Überzeugungen dieses spezifischen

Bereichs auf die akademischen Überzeugungen zu und bestimmt diese nach und nach überwiegend (horizontal gestrichelter Pfeil, Abb. 2.6, S. 76). Es gibt empirische Hinweise darauf (z.B. King, Wood & Mines, 1990), dass die epistemologischen Überzeugungen einer Person zu ihrem fachlichen Schwerpunkt im Verlauf der Hochschulausbildung gar repräsentativ für ihre akademischen epistemologischen Überzeugungen werden. Dies äußert sich darin, dass diese Personen, wenn man sie zu ihren *allgemeinen* Ansichten zu Wissen befragt, in ihren Antworten hauptsächlich durch ihre dominanten domänenspezifischen epistemologischen Überzeugungen ihres Studienbereichs beeinflusst werden. Die entsprechen wiederum mit fortschreitender Ausbildung näherungsweise den vorherrschenden Perspektiven zu Wissen des Fachbereichs, d.h. den epistemologischen Perspektiven von Fachexperten (Muis et al., 2006).

Das TIDE Framework verfügt also über drei ineinander liegende Kontextebenen. In allen drei Kontexten werden die epistemologischen Überzeugungen vom Individuum in dynamischer Interaktion mit dem Umfeld konstruiert, sind dementsprechend kontextgebunden und beeinflussen sich über die Schichten hinweg wechselseitig. Die epistemologischen Überzeugungen auf allen drei Ebenen betreffen die vier verschiedenen Dimensionen des Wissens und des Wissensprozesses, Überzeugungen zur Sicherheit, Struktur, Quelle und Begründung des Wissens (siehe graue Rauten, Abb. 2.6, S. 76). Die Spezifität bzw. der Differenzierungsgrad der individuellen Überzeugungen zum Wissen und Wissensprozess nimmt nach innen hin zu[3]. Das Modell erklärt, wie zuerst allgemeine epistemologische Überzeugungen, dann akademische und disziplinspezifische epistemologische Überzeugungen entstehen und wie sich diese gegenseitig beeinflussen. Einerseits entwickelt sich das individuelle epistemologische Denken im Laufe des Lebens mit fortschreitenden Bildungs- und Lebenserfahrungen also, indem es immer spezifischer wird, bis hin zu disziplinspezifischen Überzeugungen. Ein weiterer zweiter Entwicklungsverlauf über die Lebensspanne wird über den untersten waagerechten Pfeil in Abbildung 2.6 (S. 76) dargestellt. Diese Entwicklung des epistemologischen Denkens verläuft für alle drei Typen epistemologischer Überzeugungen (allgemeine, akademische und disziplinspezifische) innerhalb der vier Dimensionen von einer absolutistischen Perspektive, über eine multiplizistische bis hin zum Evaluatismus (Muis et al., 2006, 2016). Sie muss allerdings keineswegs linear in eine Richtung verlaufen, Rückschritte in frühere Entwicklungsstufen sind möglich, beispielsweise beim Übergang zwischen Bildungskontexten (z.B. Übergang von Grundschule auf eine weiterführende Schule) (Muis et al., 2006, S. 41). In diesem Sinne integriert dieser Aspekt des Modells die Erkenntnisse der epistemologischen Entwicklungsforschung (z.B. King & Kitchener, 1994).

Durch die Integration der verschiedensten empirischen Erkenntnisse und theoretischen Konzeptionen bietet das TIDE Framework erstens eine theoretische Basis, durch die sowohl disziplinspezifische als auch disziplinübergreifende epistemologische Überzeugungen sowie deren Zusammenspiel erklärbar und empirisch erforschbar werden. Zweitens grenzt

[3] Weitere noch spezifischere Überzeugungsebenen, wie themenspezifische epistemologische Überzeugungen oder spezielle Überzeugungen in Abhängigkeit des Wissenstyps (prozedural, deklarativ) werden vermutet, aber sind im Modell nicht integriert (Muis et al., 2006).

das TIDE Framework akademische und disziplinspezifische epistemologische Überzeugungen ab und etabliert so eine gemeinsame Sprache, die zur Beschreibung der Phänomene. Drittens erlaubt es den Vergleich von Daten über die theoretischen Ansätze hinweg und stellt viertens einen theoretischen Hintergrund zur Verfügung, der es erlaubt, die Beziehungen zwischen epistemologischen Überzeugungen und den verschiedensten Facetten der Kognition, Motivation und Leistung zu betrachten (Muis et al., 2006, S. 30).

Zusammenfassend lässt sich bezüglich der Frage der Disziplinspezifität epistemologischer Überzeugungen folgendes festhalten: Es wird aufgrund der vielen Forschungsergebnisse mittlerweile weitgehend davon ausgegangen, dass Individuen sowohl über disziplinspezifische als auch über disziplinübergreifende epistemologische Überzeugungen verfügen (Hofer, 2016). Darüber hinaus herrschen in den verschiedenen akademischen Disziplinen unterschiedliche Überzeugungen zu Wissen und dem Wissensprozess vor, die von Expertinnen und Experten der jeweiligen Disziplin vertreten werden. Akademische Disziplinen unterscheiden sich bezüglich der Standards der Beweisführung und Begründung von Wissen, der Struktur des Wissens und der Wahrscheinlichkeit sicheren Wissens. In den unterschiedlichen Disziplinen werden jeweils andere methodologische Ansätze und Arten der Begründung von Wissensprozessen genutzt (Hofer, 2006b, S. 68). Mit der akademischen Ausbildung nähern sich die Lernenden (zunächst Laien in einer Disziplin) immer weiter den epistemologischen Überzeugungen der Lehrenden (Expertinnen und Experten der jeweiligen Disziplin) an (Muis et al., 2006).

## 2.7 Verortung epistemologischer Überzeugungen in Theorien der Kognition und Metakognition

Seit dem Beginn der Forschung zu epistemologischer Entwicklung wurden die epistemologischen Überzeugungen auf einer Metaebene des Denkens gesehen und so mit der Metakognition in Verbindung gebracht (Perry 1970, 1999; Kitchener, 1983). Um das Konstrukt epistemologischer Überzeugungen noch genauer einzugrenzen und zu definieren, wird im Folgenden dargestellt, wie sich die epistemologischen Überzeugungen in der Kognition und Metakognition verorten lassen.

Flavell legte mit seiner Forschung über das Wissen und Nachdenken von Menschen über ihre eigenen kognitiven Phänomene (Flavell, 1971, 1979; Flavell & Wellman, 1977) den Grundstein für die Forschung über Metakognition (Veenman, 2012, S. 21). Metakognition ist Kognition auf der Metaebene und bezeichnet das Wissen und Denken eines Individuums über die eigenen kognitiven Prozesse. Sie dient dem Menschen dazu, kognitive Aufgaben, wie Problemlösen und Lernen, zu meistern (Flavell, 1979).

In ihrem bedeutenden Metakognitionsmodell gehen Nelson und Narens (1994) davon aus, dass kognitive Prozesse auf mindestens zwei Ebenen operieren, auf der Objektebene und der Metaebene. Der Objektebene enthält die Repräsentation der kognitiven Aufgabe, wie beispielsweise einer Lernsituation (Lernumgebung, Lerninhalt, Bedingungen des Lerners), auf die sich die Metakognition bezieht. Der Metaebene ist die Ebene der Metakognition. Er enthält ein unvollständiges, dynamisches Modell darüber, was auf dem Objektebene passiert (Nelson & Narens, 1994). Kognition und Metakognition interagieren demnach miteinander. Die Metakognition greift auf die Kognition zurück und beeinflusst diese wiederum (Veenman, van Hout-Wolters & Afflerbach, 2006).

Metakognition besteht aus den Komponenten *Wissen über Kognition* (metakognitives Wissen) und *Regulation der Kognition* (metakognitive Fähigkeiten) (Schraw, Olafson, Weibel & Sewing, 2012). Das metakognitive Wissen umfasst das *deklarative* (wissen was), *prozedurale* (wissen wie) und *konditionale Wissen* (wissen wann und warum) eines Individuums über die eigene Kognition. Das metakognitive deklarative Wissen ist solches, das ein Individuum über sich selbst als Lerner hat und darüber, welche Faktoren die eigene Lernleistung beeinflussen (wie z.B. Einschränkungen des Gedächtnisses: Was weiß ich, was nicht?). Metakognitives prozedurales Wissen bezieht sich auf das verfügbare Repertoire an Lernstrategien (z.B. Wiederholungsstrategien, Elaborationsstrategien, Organisationsstrategien). Metakognitives konditionales Wissen ist das Wissen eines Individuums darüber, welche Lernstrategie wann und warum eingesetzt werden muss. Es befähigt den Lerner, die Anforderungen einer Lernsituation richtig einzuschätzen und entsprechend der Umstände eine angemessene Strategie auszuwählen (Schraw et al., 2012). Der zweite Bestandteil der Metakognition, die Regulation der Kognition, bezieht sich auf die metakognitiven Fähigkeiten, eigene Lern- und Denkprozesse selbstreguliert zu planen, zu überwachen und zu evaluieren sowie Informationsorganisationsstrategien und Fehlerbehebungsstrategien. Für die Planung der Lern- und Denkprozesse wählt das Individuum angemessene Strategien, setzt

sich Ziele, aktiviert relevantes Hintergrundwissen und teilt sich Zeit ein und stellt ganz allgemein verfügbare kognitive Ressourcen bereit. Um den eigenen Lernprozess zu steuern, muss er ständig überwacht werden. Dazu muss der Lerner über die Fähigkeit verfügen, sich selbst zu überprüfen (*self-testing*). Der Lerner muss sowohl seinen Lernprozess als auch die Selbstregulation seines Lernens evaluieren. Das beinhaltet zum Beispiel das erneute Bewerten der gesetzten Ziele, das Anpassen von Vorhersagen und das Festigen des Gelernten. Während des Lernens können Informationsorganisationsstrategien eingesetzt werden, um die Effizienz der Informationsverarbeitung zu steigern. Mittels der Fehlerbehebungsstrategien werden während des Lernens Verständnisfehler und Performanzfehler korrigiert (vgl. Schraw et al., 2012, S. 58 f.).

Seit dem Beginn der Forschung zu epistemologischer Entwicklung wurde epistemologisches Denken als Denken auf einer Metaebene betrachtet und mit der metakognitiven Entwicklung assoziiert. Schon Perry (1970) beschrieb den intellektuellen Entwicklungsschritt, wenn die Studierenden in seiner Studie herausfanden „how to think further, how to think relatively and contingently, and how to think about thinking“ (S. 41) als den wichtigsten Schritt überhaupt. Er beschrieb die Transition vom dualistischen Denken zum relativistischen Denken als Transition vom Denken zum Meta-Denken, im Sinne einer Transition von einer denkenden Person zu einer Person, die Kritiker ihres eigenen Denkens ist (Perry, 1970). Meta-Denken ist nach Perry essentiell für die epistemologische Entwicklung, da relativistisches Denken, nach seiner Konzeption, mit einschließt, sich der eigenen Denkweisen bewusst zu sein, sie mit anderen Arten des Denkens zu vergleichen und verschiedene Ansätze abzuwägen (ebd.). Epistemologisches Denken wurde seit Perrys Arbeiten in verschiedensten Forschungssträngen als metakognitives Konstrukt aufgefasst (z.B. Barzilai & Zohar, 2016; Bromme, 2010; Hofer, 2004a, 2016; Kitchener, 1983; Kuhn, 1999; Mason, Boldrin & Ariasi, 2010a; Moshman, 2015; Richter & Schmid, 2010; Schommer-Aikins, 2004).

Kitchener (1983) stellt explizit, basierend auf den Arbeiten von Flavell (Flavell, 1971, 1979; Flavell & Wellman, 1977), den Zusammenhang zwischen Metakognition mit den epistemologischen Überzeugungen her (*epistemic cognition*, Kitchener, 1983), indem sie sie als *meta-knowing* – Wissen über Wissen – konzeptualisiert (vgl. Hofer, 2004a, S. 46). Sie verortet die individuellen epistemologischen Überzeugungen, die sie *epistemic cognition* nennt, in Bezug zu Kognition und Metakognition auf der höchsten von drei hierarchischen Ebenen der kognitiven Verarbeitung (Kitchener, 1983). Ihr *3-Level model of cognitive processing* besteht aus drei aufeinander aufbauenden Ebenen der kognitiven Verarbeitung: Kognition, Metakognition und epistemische Kognition (siehe Tab. 2.3, S. 82). Auf der Ebene der Kognition finden Prozesse des Wahrnehmens, Lesens und Berechnens statt. Die Ebene der Metakognition betrifft das Wissen über kognitive Aufgaben wie Strategienutzung und Überwachung. Die höchste Ebene ist die der epistemischen Kognition, die noch über der Metakognition angesiedelt ist und in Verbindung mit ihr aber auch der Kognition die Überwachung der epistemischen Beschaffenheit des Problemlösens ermöglicht, inklusive des Wahrnehmens von Grenzen des Wissens und Sicherheit des Wissens und Kriterien, die im Prozess des Wissenserwerbs eine Rolle spielen (Hofer, 2004a).

TABELLE 2.3: Verortung des epistemologischen Verstehens in der kognitiven Verarbeitung bei Kitchener (1983) und Kuhn und Pearsall (2000) nach Hofer (2001, S. 364), in eigener Darstellung

| 3-Ebenen-Modell der kognitiven Verarbeitung (Kitchener, 1983) | 3-Ebenen-Modell des Meta-Wissens (Kuhn & Pearsall, 2000) |
|---|---|
| KOGNITION<br>Wahrnehmen, Lesen, Berechnen | |
| METAKOGNITION<br>Wissen über kognitive Aufgaben | METAKOGNITIVES WISSEN<br>Wissen über deklaratives Wissen<br>METASTRATEGISCHES WISSEN<br>Wissen über prozedurales Wissen |
| EPISTEMISCHE KOGNITION<br>Überwachung der epistemischen Beschaffenheit des Problemlösens | EPISTEMOLOGISCHES WISSEN<br>Wissen über Wissen und den Prozess des Wissens (eigenes und fremdes) |

Kuhns (2000) *3-Level model of meta-knowing* ist ähnlich dem von Kitchener aufgebaut. Auch sie sieht das epistemologische Verstehen auf der höchsten von drei hierarchischen Ebenen angesiedelt. Das *meta-knowing* als Kognition, die Kognition selbst zum Objekt hat, hat drei Ebenen: *metacognitive knowing*, *metastrategic knowing* und *epistemological meta-knowing*. Die erste Ebene umfasst Wissen über deklaratives Wissen (Wissen als Produkt). Die zweite betrifft Wissen über prozedurales Wissen (Wissen als Prozess), während die dritte den abstrakteren Prozess des Wissens über das eigene oder fremde Wissen und den Prozess des Wissens betrifft (siehe Tab. 2.3). Kuhn konzeptualisiert epistemologisches Denken somit als metakognitives Wissen über die allgemeine Beschaffenheit des Wissens und den Wissensprozess. Ein Beispiel für solches Wissen ist das Wissen eines Individuums über die eigene Rolle in der Wissenskonstruktion und über die Beschaffenheit von Wahrheit und Wissensbegründung (Kuhn, 1999; Moshman, 2015).

Gemeinsam ist den beiden Modellen von Kuhn (2000) und Kitchener (1983), dass sie die epistemologischen Überzeugungen als einzelne, spezifische metakognitive Komponente betrachten. In beiden Modellen wird epistemologisches Denken auf höchster Ebene und als eine fortgeschrittene Form der Metakognition beschrieben. Kuhn und Kitchener gehen übereinstimmend davon aus, dass sich die jeweils dritte Ebene erst in der späten Adoleszenz entwickelt. Erforscht wurde das epistemologische Verstehen auf metakognitiver Ebene über das Lösen komplexer, schlecht strukturierter Probleme (King & Kitchener, 1994), das Argumentieren (Kuhn, 1991) und das Lösen komplexer Aufgaben, wie Schlussfolgerungsprozesse von Geschworenen (Kuhn & Pearsall, 2000).

Als übergreifende Verbindung zwischen den Modellen von Kitchener (1983) und Kuhn und Pearsall (2000) entwickelte Hofer (2004a) ihre Konzeption epistemologischer Überzeugungen (Hofer & Pintrich, 1997, siehe Kapitel 2.2.5, S. 43) weiter als *epistemic metacognition* (Hofer, 2004a). Ziel von Hofers (2004a) Weiterentwicklung ist es, das Konstrukt epistemologischer Überzeugungen klarer zu fassen und besser einzugrenzen, indem sie es im Konstrukt der Metakognition als sogenannte *epistemic metacognition* verortet. In Abbildung 2.7 ist diese Weiterentwicklung dargestellt.

| Metakognition - existierende Komponenten | Metakognition - epistemische Komponenten |
|---|---|
| Metakognitives Wissen | Überzeugungen über die *Natur des Wissens* |
| Wissen über Kognition und Strategien, Aufgaben und Kontexte, sich selbst als Lerner und denkende Person | *Sicherheit des Wissens:* Implizite Uberzeugungen darüber, ob Wissen sicher oder vorläufig und sich entwickelnd ist<br>*Einfachheit des Wissens:* Implizite Überzeugungen darüber, ob Wissen aus einer einfachen Ansammlung von Fakten besteht oder komplex und untereinander verbunden ist. |
| | Überzeugungen über sich selbst als etwas wissende Person (Self as Knower) |
| | Epistemologische Disposition, *need for cognition*, *need for epistemic closure* |
| Metakognitive Bewertung und Überwachung | Überzeugungen über die *Natur des Wissensprozesses* |
| Überwachung des Verstehens und Lernens, Einschätzen der Aufgabenschwierigkeit (z.B.: "Verstehe ich das?") | Bewertung der *Quelle des Wissens:* "Woher weiß ich das?", "Wie passt dies zu meiner eigenen Erfahrung?" |
| | Festlegen der *Begründung des Wissens:* "Schätze ich dies als zuverlässig ein?", "Liegen Beweise vor, die diese Aussage stützen?", "Lassen sich Theorie und Evidenzen vereinbaren?" |
| Selbstregulation und Kontrolle der Kognition (z.B.: Muss ich das erneut lesen?") | *Regulieren der Kognition während der Wissenskonstruktion* |
| | Selbstregulation des Wissensprozesses, beeinflusst von Volition, Interesse, Motivation, Dispositionen des Denkens, intellektuellen Werten und Überzeugungen: "Weiß ich, was ich wissen muss oder muss ich mehr wissen?", "Wie gehe ich das an?" |

ABBILDUNG 2.7: Epistemologische Prozesse (*epistemic metacognition*) integriert als Erweiterung der Metakognition nach Hofer (2004a, S. 49), in eigener Darstellung

Epistemologisches Denken umfasst bei Hofer sowohl metakognitives Wissen über die Natur des Wissens und den Wissensprozess als auch metakognitive Fertigkeiten, wie Überwachung und Kontrolle der Wissenskonstruktion und Evaluation (Hofer, 2004a). Das epistemologische Denken stellt in diesem Modell demnach keine einzelne metakognitive Komponente mehr dar, vielmehr fügen sich die einzelnen Aspekte des epistemologischen Denkens in die verschiedenen metakognitiven Kategorien ein. Überzeugungen zur Natur des Wissens, die die Dimensionen *Sicherheit* und *Einfachheit des Wissens* umfasst sowie Überzeugungen über sich selbst als wissende Person, gliedert sie in die Komponente *metakognitives Wissen*. Die Überzeugungen zur Natur des Wissensprozesses mit den Dimensionen *Quelle* und *Begründung des Wissens* gehören zur metakognitiven Komponente der Bewertung und Überwachung des eigenen Verstehens und Lernens (siehe Abb. 2.7, S. 83). Die ursprünglichen metakognitiven Komponenten (siehe linke Spalte, Abb. 2.7, S. 83) werden also breiter gefasst, um die epistemologischen Elemente konzeptuell einzuschließen (siehe rechte Spalte, Abb. 2.7, S. 83). Hofer (2004a) betrachtet epistemologische Überzeugungen nun als

Aspekt des metakognitiven Bewusstseins (*metacognitive awareness*), die nicht nur beim Bearbeiten schlecht strukturierter Probleme (King & Kitchener, 1994) und Argumentieren über komplexe Probleme (Kuhn, 1991) aktiviert würden, sondern darüber hinaus auch bei alltäglichen Prozessen der Wissensaneignung und -konstruktion (Hofer, 2004a, S. 44). Da die Metakognition auf die Kognition zugreift und diese wiederum speist (Veenman et al., 2006), d.h. die Kognition und Metakognition miteinander interagieren (Nelson & Narens, 1994), ist nachvollziehbar, weshalb Hofer (2004a) nun annimmt, dass die epistemologischen Überzeugungen sowohl auf metakognitiver als auch auf kognitiver Ebene operieren, auch wenn sie auf der metakognitiven verortet werden.

Epistemologische Überzeugungen werden nach wie vor als System von Überzeugungsdimensionen betrachtet, die ähnlich zu Theorien organisiert sind. Diese subjektiven Theorien über Wissen und den Wissensprozess fungieren auch als Ziele und haben somit auch motivationale Konsequenzen für den Lernprozess (Hofer, 2004a). Dies kann an einem Beispiel illustriert werden: Eine Person, die der Überzeugung ist, dass Wissen aus abgeschlossenen, einzelnen Fakten besteht und es nur eine richtige Antwort (einen speziellen Fakt) auf eine Frage geben kann, wird konkurrierende Informationen zu einem Sachverhalt (die in einer oder auch mehreren Wissensquellen dargestellt werden) als verwirrend und unnötig empfinden. Bei einer Recherche nach Antworten auf eine gestellte Frage wird diese Person aufgrund ihrer epistemologischen Überzeugungen also gar nicht nach verschiedenen möglichen Antworten suchen, sondern nur nach der vermeintlich einen richtigen. Dies kann darin resultieren, dass die Person nach Auffinden einer ersten Antwort die Recherche abbricht, in dem Glauben, die *eine* Antwort gefunden zu haben. Eine Suche nach konkurrierenden Antworten und Perspektiven auf ein Thema wird nicht in Betracht gezogen (Zielsetzung). Trifft die Person doch auf konkurrierende Informationen, wird sie keinen Versuch unternehmen diese zu bewerten und zu gewichten, da sie annimmt, es gebe nur eine eindeutige Wahrheit (Hofer, 2004a, S. 47). Epistemologische Überzeugungen entstehen in Interaktion mit der Umwelt, werden von Kultur und Bildung und anderen Kontextvariablen beeinflusst und operieren sowohl auf disziplinübergreifender Ebene als auch auf disziplinspezifischer Ebene (ebd.). Sie kommen bei der Begegnung mit neuen Informationen und im Kontext der Wissenskonstruktion zum Einsatz und zwar immer, wenn miteinander konkurrierende Wissensansprüche die Aufmerksamkeit und Auflösung erfordern – auch, aber nicht nur beim Lösen komplexer, schlecht strukturierter Probleme. Sie sind also in der Anwendung situiert und werden im Kontext aktiviert Hofer (2001).

Aus den vorangegangenen Darstellungen geht hervor, dass epistemologische Überzeugungen über die verschiedensten Konzeptionen hinweg als metakognitives Konstrukt aufgefasst werden (z.B. Barzilai & Zohar, 2016; Bromme, 2010; Hofer, 2004a, 2016; Kuhn & Pearsall, 2000; Mason et al., 2010a; Moshman, 2015; Richter & Schmid, 2010; Schommer-Aikins, 2004), die sowohl auf kognitiver Ebene als auch auf metakognitiver Ebene operieren (Barzilai & Zohar, 2012; Hofer, 2005, 2016; Richter & Schmid, 2010). In dieser Arbeit wird damit übereinstimmend angenommen, dass sich epistemologische Überzeugungen, wie oben dargestellt (siehe Abb. 2.7, S. 83), integriert in den Komponenten der Metakognition verorten lassen.

## 2.8 Der Einfluss epistemologischer Überzeugungen auf den Lehr-Lernprozess

Epistemologische Überzeugungen nehmen auf verschiedene Weise Einfluss auf den Lehr-Lernprozess (Hofer, 2001; Schommer-Aikins, 2004). Das umfasst sowohl akademische Lehr-Lernprozesse, wie sie in Schule und Universität stattfinden, als alle informellen Lehr-Lernprozesse. Daraus ergibt sich die Relevanz dieses Konstrukts für das Lehren und lebenslange Lernen (Hofer, 2001). In diesem Kapitel werden die Einflüsse epistemologischer Überzeugungen auf das Lernen von Individuen sowie das Lehren dargestellt. Dazu wird zu Beginn kurz dargestellt, welchen Fragen des Lehrens und Lernens die Pädagogische Psychologie, die Entwicklungspsychologie und die Fachdidaktiken bezüglich der epistemologischen Überzeugungen hauptsächlich nachgegangen sind. Als Basis für die Darstellung der Beziehungen epistemologischer Überzeugungen zum Lehr-Lernprozess dienen das Modell von Hofer (2001) und das *Embedded Systemic Model* von Schommer-Aikins (2004), die in der Forschungstradition der pädagogischen Psychologie anzusiedeln sind. Es werden die verschiedensten empirischen Erkenntnisse zum Einfluss epistemologischer Überzeugungen auf den Lehr-Lernprozess vorgestellt. Ein Schwerpunkt wird die Beziehung zwischen epistemologischen Überzeugungen und Lernstrategien sein, da diese für die vorliegende Arbeit von besonderem Interesse sind. Aus diesem Grund wird im Rahmen dieses Kapitels auch auf die Lernstrategien im Studium eingegangen.

Die verschiedenen Forschungsdisziplinen widmen sich den epistemologischen Überzeugungen von Lernenden und Lehrenden. Empirische Studien dazu finden sich in den Forschungsdisziplinen der Psychologie und pädagogischen Psychologie (vgl. z.B. Hofer, 2001; Schmid & Lutz, 2007; Schommer-Aikins, 2003), der Entwicklungspsychologie (vgl. z.B. King & Kitchener, 1994; Kuhn & Weinstock, 2002) sowie der Erziehungswissenschaft und in den fachdidaktischen Forschungsbereichen (vgl. z.B. Schoenfeld, 1983, 1985, 1998; Wineburg, 1991). In der entwicklungspsychologischen Forschung zu epistemologischen Überzeugungen wird den individuellen epistemologischen Überzeugungen, die dem Problemlösen und Argumentieren im Alltag oder bei akademischen Fragen zugrundeliegen sowie deren Einfluss darauf nachgegangen (King & Kitchener, 1994; Kuhn, 1991). Die Rechtfertigung und Begründung von Wissensannahmen und -urteilen steht hier im Vordergrund (Zinn, 2013, S. 78). In pädagogisch-psychologischer Forschungstradition wird die Beziehung epistemologischer Überzeugungen zum Lehr-Lernprozess im akademischen Kontext vor allem bei Studierenden oder Lehrenden untersucht. Die fachdidaktischen Forschungsarbeiten untersuchen ebenfalls die Rolle der epistemologischen Überzeugungen beim Lernen und Lehren, allerdings beziehen sich die Studien hier vielfach auf Schülerinnen und Schüler sowie Lehrkräfte. Die fachdidaktischen Studien fokussieren sich außerdem hauptsächlich auf die disziplinspezifischen epistemologischen Überzeugungen, d.h. die Überzeugungen zum Wissen in der Fachdisziplin (z.B. Geschichte, Mathematik). Fachdidaktikerinnen und -didaktiker interessieren sich u.a. auch für Überzeugungen von Lehrenden und Lernenden, die spezifisch für das Fach sind. Hier wird beispielsweise der Frage nachgegangen, ob in Mathematik

spezielle Arten und Weisen vorherrschen, wie Wissen begründet und widerlegt wird. Vor allem in den Disziplinen Mathematik (z.B. de Corte, Op't Eynde & Verschaffel, 2002; Köller et al., 2000; Muis, 2004; Schoenfeld, 1992, 1998; Schommer-Aikins, 2008) und den Naturwissenschaften[4] entstanden starke, eigene Forschungslinien mit zahlreichen Publikationen, die sich mit den individuellen epistemologischen Überzeugungen bezüglich des Fachwissens befassen (in den Naturwissenschaften z.B. Bell & Linn, 2002; Elder, 2002; Qian & Alvermann, 1995; Qian & Pan, 2002; Sandoval, 2005; Sinatra & Chinn, 2012 und im Speziellen in der Physik z.B. Conley et al., 2004; Hammer & Elby, 2002, 2003; Louca et al., 2004; Stahl & Bromme, 2007; Urhahne & Harms, 2006; Urhahne & Hopf, 2004). Von den Fachdisziplinen, die an Schule und Universität gelehrt werden, wurde die Relevanz des Konstrukts epistemologischer Überzeugungen für den Lehr-Lernprozess in Mathematik und den Naturwissenschaften besonders früh realisiert. Ein etwas weniger forschungsstarkes Fach ist Geschichte (z.B. Maggioni et al., 2010; Wineburg, 1991). Auch zum Bereich *Literary Reasoning* (z.B. Richter, 2003, 2011; Bråten & Strømsø, 2006; Garner & Alexander, 1994; Strømsø & Bråten, 2011a; Strømsø & Kammerer, 2016) entwickelte sich ein eigener Forschungsstrang, der die Rolle der individuellen epistemologischen Überzeugungen beim Umgang, Lesen, Verstehen und Schlussfolgern aus Texten betrifft.

Für eine aktuelle, überblicksartige Darstellung der Forschungsstränge der genannten Disziplinen sei auf betreffenden Beiträge[5] im Sammelband von Greene et al. (2016) verwiesen (siehe auch Sandoval, 2016). Ein weiterer Bereich, der je nach Forschungsschwerpunkt in der pädagogisch-psychologischen Forschung oder aber der Fachdidaktik anzusiedeln ist, sind die *teacher beliefs*. Hier entwickelte sich ein eigener Forschungsstrang, der sich einerseits mit den epistemologischen Überzeugungen von Lehrkräften (z.B. Bendixen, 2016; Brownlee et al., 2016; Buehl & Fives, 2016; Chan & Elliott, 2002; Elby & Hammer, 2010; Fives & Buehl, 2010, 2012; Muis & Foy, 2010; Olafson & Schraw, 2010; Schraw, Olafson & VanderVeldt, 2011; Tabak & Weinstock, 2011; Valanides & Angeli, 2011; Weinstock & Roth, 2011) und der Rolle von epistemologischen Überzeugungen in der Lehrerbildung befasst (z.B. Bendixen & Corkill, 2011; Brownlee, Ferguson & Ryan, 2017; Fives, 2011; Marra & Palmer, 2011; Sinatra & Kardash, 2004; Strømsø & Bråten, 2011a; Tillma, 2011; Walker, Brownlee, Exley, Woods & Whiteford, 2011; Yadav, Herron & Samarapungavan, 2011). Gleichzeitig umfasst der Begriff *teacher beliefs* und die Forschung dazu, neben den epistemologischen Überzeugungen, noch viel weiter gefasst auch die Überzeugungen von Lehrkräften zu u.a. Lernen, Lehren, Lernenden, Lehrenden, der gesellschaftlichen Rolle der Schule und des Curriculums (Levitt, 2001) und untersucht die komplexe Beziehung zwischen diesen Überzeugungen mit dem Lehrerhandeln (vgl. z.B. Abrami, Poulsen & Chambers, 2004) sowie der Lernleistung von Schülerinnen und Schülern (vgl. z.B. Muis & Franco, 2010; Pajares, 1992). Einen guten Überblick darüber, was unter den *teachers beliefs* verstanden wird, welche identifiziert wurden und was darüber herausgefunden wurde, geben Fives und Buehl (2012).

[4] Die Naturwissenschaften umfassen die akademischen Disziplinen Physik, Biologie und Chemie.

[5] Mathematik (Depaepe et al., 2016), Naturwissenschaften (Elby et al., 2016), Geschichte (VanSledright & Maggioni, 2016), Literary Reasoning (Lee et al., 2016)

### 2.8.1 Modelle zum Einfluss epistemologischer Überzeugungen auf den Lehr-Lernprozess

Das Wirkmodell von Hofer (2001) nimmt einen indirekten Einfluss der individuellen epistemologischen Überzeugungen auf den Lernerfolg des Individuums an. „Learners arrive in the classroom with existing epistemological beliefs and theories that lead to interpretations of instruction, and, as these beliefs change, so do these interpretations." (Hofer, 2001, S. 372). Die epistemologischen Überzeugungen der Lernenden beeinflussen deren Überzeugungen zu Lernen und Unterricht sowie deren Lernmotivation und Wahl der Lernstrategien (siehe Abb. 2.8). Die Überzeugungen zum Lernen und Unterricht wiederum sowie die Lernmotivation beeinflussen die Wahl der Lernstrategien. Die Lernstrategien und Lernmotivation wirken sich ihrerseits wieder auf den Lernprozess aus. Auch die Lehrenden verfügen jedoch über epistemologische Überzeugungen, die sich auf ihr pädagogisches Handeln auswirken sowie darauf, welche Aufgaben sie sich selbst als Lehrende im Unterricht zuschreiben. Darüber werden die epistemologischen Überzeugungen der Lernenden beeinflusst (siehe Abb. 2.8). Zwar bezog Hofer (2001) dieses Modell ursprünglich auf das Lernen in der Schule, allerdings kann es auch auf andere akademische Lernprozesse übertragen werden, wie beispielsweise auf Lernprozesse bei Erwachsenen im universitären Kontext (Urhahne & Harms, 2006).

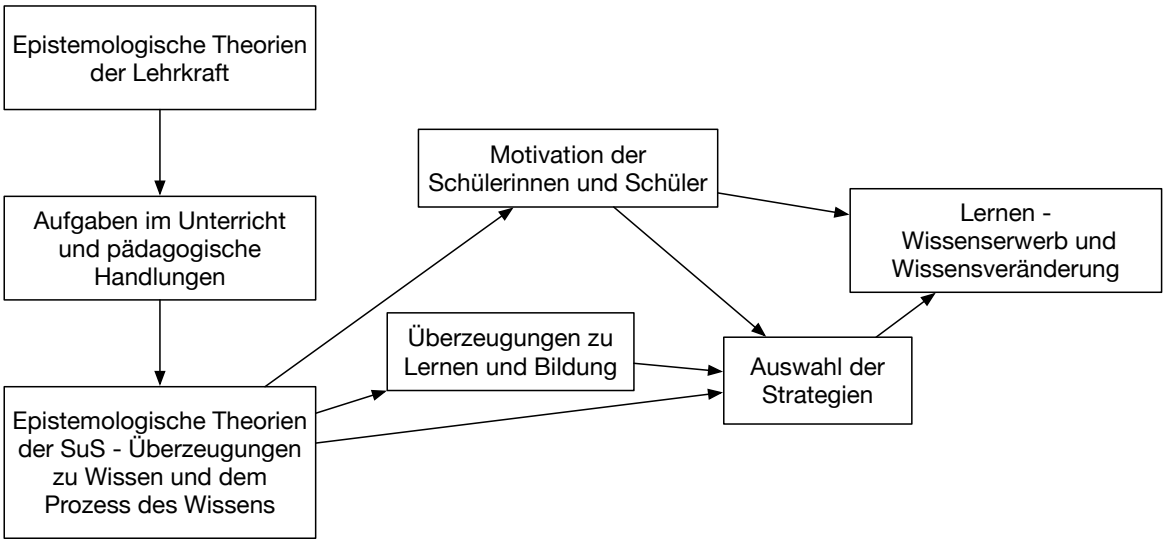


ABBILDUNG 2.8: Einfluss epistemologischer Überzeugungen auf den Lernprozess nach Hofer (2001, S. 372), in eigener Darstellung

Schommers *Embedded Systemic Model* (Schommer-Aikins, 2004) stellt epistemologische Überzeugungen im Zusammenspiel mit kulturellen Sichtweisen, den sogenannten *Ways of Knowing* (*connected knowing*, *separate knowing*) und der Selbstregulation und Leistung von Lernern dar (siehe Abb. 2.9, S. 88). Im Unterschied zu ihrer früheren Konzeptualisierung von 1990 differenziert Schommer nun explizit zwischen *Überzeugungen zu Wissen* und *Überzeugungen zu Lernen* (Schommer-Aikins, 2004, S. 25). Zwar gehören die *Überzeugungen zu Lernen* nach wie vor zum Konstrukt epistemologischer Überzeugungen, allerdings wird der

Einfluss der Überzeugungen zu Wissen und Lernen auf andere Variablen, wie Lernleistung und selbstreguliertes Lernen, nun im Rahmen des neuen Modells getrennt betrachtet. Die beiden Dimensionen *speed of learning* und *ability to learn* fasst Schommer in der übergeordneten Dimension *beliefs about learning* zusammen. Die drei Dimensionen *source of knowledge*, *structure of knowledge* und *justification of knowledge* werden dagegen der übergeordneten Dimension *beliefs about knowledge* zugeordnet (siehe Abb. 2.9). Hier zeigt sich, dass Schommer bei den Überzeugungen zum Wissen nicht mehr von ihren ursprünglichen Dimensionen *certain knowledge* und *simple knowledge* spricht (vgl. angenommene Struktur des SEQ, Tab. 2.2, S. 40). Statt von einfachem vs. komplexem Wissen wird nun von der Dimension *Struktur des Wissens* gesprochen. Die Dimension *justification of knowledge* kommt bei Schommer ursprünglich gar nicht vor, sondern wurde von Hofer und Pintrich (1997) als eine der Kerndimensionen epistemologischer Überzeugungen modellübergreifend identifiziert (siehe Schommer-Aikins, 2004, S. 25).

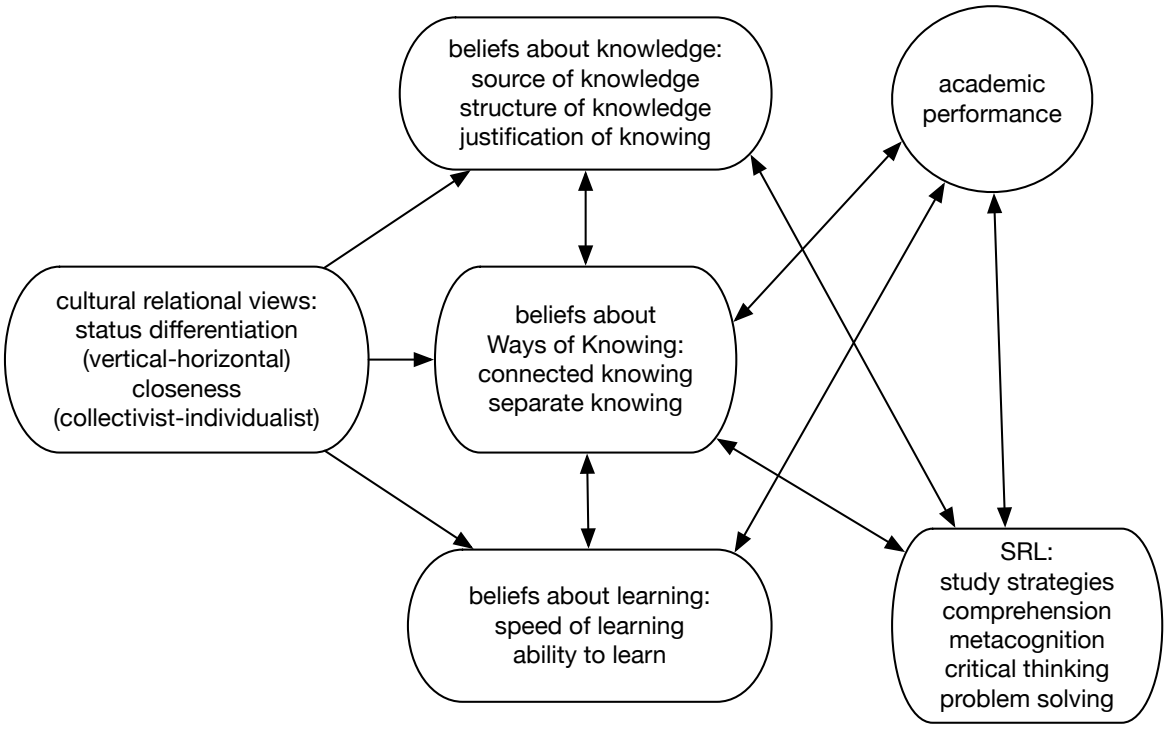


ABBILDUNG 2.9: *Embedded Systemic Model* nach Schommer-Aikins (2004, S. 25), in eigener Darstellung

Die *Ways of Knowing* beziehen sich auf Überzeugungen zum Wissenserwerb. Schommer-Aikins nimmt an, dass diese Überzeugungen zum *connected knowing* und *separate knowing* das selbstregulierte Lernen und die Lernleistung beeinflussen (vgl. Schommer-Aikins, 2004; Schommer-Aikins et al., 2010). Das Konstrukt stammt von Galloti, McVicker Clinchy, Ainsworth, Lavin und Manfield (1999), die die *Women's Ways of Knowing* von Belenky et al. (1986) weiterentwickelten, indem sie ein Instrument zur Erfassung von *connencted knowing* und *separate knowing* entwickelten. Lerner mit starken Überzeugungen zur Verbundenheit des Wissensprozesses gehen den Wissenserwerb insofern anders an, als dass sie sich zunächst

über einen Perspektivwechsel in die dargestellte Sichtweise hineinversetzen, um den zentralen Punkt nachvollziehen zu können und sich im Anschluss kritisch damit auseinanderzusetzen. Lerner mit starken Überzeugungen zur Getrenntheit des Wissensprozesses versetzen sich nicht in die dargestellte Sichtweise, sondern betrachten diese aus entgegengesetzter Perspektive kritisch. Sie warten zunächst auf Belege für die dargestellte Sichtweise, bevor sie bereit sind, sich auf tiefe Auseinandersetzung damit einzulassen (Schommer-Aikins, 2004). Schommer-Aikins und Easter (2006) stellten fest, dass beide *Ways of Knowing*, mediiert über die Überzeugung zur Lerngeschwindigkeit (*speed of learning*), die akademische Leistung prädizieren. Je mehr die Lerner von der Verbundenheit und Getrenntheit des Wissenserwerbs überzeugt waren, desto mehr waren sie auch davon überzeugt, dass Lernen graduell über Zeit erfolgt. Je mehr sie Lernen wiederum für graduell hielten, desto bessere Leistungen erbrachten sie in einem Kurs, der Lesen, Schreiben und Präsentieren erforderte (Schommer-Aikins & Easter, 2006).

Schommer-Aikins (2004) nimmt an, dass kulturelle Einflüsse auf die Überzeugungen zu Wissen, Wissensprozess und Lernen wirken und über diese Pfade auch auf das selbstgesteuerte Lernen und die Lernleistung. Unter *cultural relational views* versteht sie die kulturell geprägten Sichweisen von Personen zu sozialen Beziehungen. Diese können sich im angenommenen Grad der Nähe (*collectivist* vs. *individualist*) und in wahrgenommenen Statusunterschieden (*vertical* vs. *horizonal*) unterscheiden. Der Grad der Nähe bezieht sich darauf, inwiefern Personen Abhängigkeit (*individualist*) oder Unabhängigkeit (*collectivist*) von anderen Personen annehmen. Personen mit kollektivistischem kulturellen Hintergrund sehen sich als Teil des Kollektivs mit den anderen Mitgliedern eng verbunden. Kollektive Ziele und Bedarfe werden als wichtiger erachtet als die individuellen. In Kulturen mit individualistischer Orientierung steht das Individuum mit seinen Bedürfnissen und Zielen im Zentrum. Die wahrgenommenen Statusunterschiede beziehen sich darauf, ob Personen auf Basis ihres kulturellen Hintergrundes an den gleichen Status aller Personen (*horizonal*) oder aber an hierarchische Strukturen (*vertical*) mit starken Status- und Machtunterschieden glauben. Auch Personen mit dem gleichen kulturellen Hintergrund können Statusunterschiede in unterschiedlicher Ausprägungen horizonal oder vertikal betrachten. Diese Ausprägungen sind relevant für das Lernen, da sie mit den epistemologischen Überzeugungen zusammenhängen. Lerner mit wenig fortgeschrittenen Überzeugungen sehen ihr Verhältnis zu Experten/Autoritäten (z.B. der Lehrkraft) als stark hierarchisch. Sie sehen sich als passive Empfänger von Wissen, das ihnen von einer allwissenden Autorität vermittelt werden muss. Epistemologisch fortgeschrittenere Lerner sind der Überzeugung, dass Wissen aus empirischer Evidenz und Logik resultieren, derer sie selbst nachgehen können und betrachten ihre Beziehung zu Experten ausgeglichener (Belenky et al., 1986; Kitchener et al., 1989). Auch die Bereitschaft von Lernern, Fragen zu stellen und Kritik zu akzeptieren hängt, nach der Forschung Baxter Magoldas (Baxter Magolda, 1992) mit der empfundenen Nähe zum Lehrenden und den epistemologischen Überzeugungen der Lerner zusammen (vgl. Schommer-Aikins, 2004, S. 25).

Zusammengefasst geht Schommer-Aikins (2004) also davon aus, dass sich die epistemologischen Überzeugungen zum Wissen und Wissenserwerb auf die Lernstrategien, das Verstehen, die Metakognition, das kritisches Denken und das Problemlösen auswirken und so das selbstregulierte Lernen als Ganzes beeinflussen. Auch im Modell von Hofer (2001) findet sich der direkte Pfad zwischen epistemologischen Überzeugungen und Lernstrategien. In beiden Modellen mediiert die Auswahl der Lernstrategien den Einfluss, den die epistemologischen Überzeugungen auf die Lernleistung haben.

### 2.8.2 Epistemologische Überzeugungen beim Lernen

Die Beziehung zwischen den individuellen Überzeugungen zu Wissen und dem Wissensprozess mit dem Lernprozess konnte vielfach empirisch bestätigt werden (z.B. Buehl & Alexander, 2004; Kardash & Noel, 2000; M. P. Ryan, 1984; Schommer, 1990; Schommer et al., 1992). Im Folgenden werden einige Befunde vorgestellt, die sich auf die in den oben beschriebenen Modellen (Hofer, 2001; Schommer-Aikins, 2004) angenommenen Beziehungen zwischen epistemologischen Überzeugungen und den verschiedensten Aspekten des Lernprozesses beziehen, wie beispielsweise die akademische Leistung, das fachspezifische Selbstkonzept, die Motivation und Lernstrategien. Der Schwerpunkt liegt dabei auf der Verbindung zwischen den Kerndimensionen epistemologischer Überzeugungen (Hofer, 2001; Hofer & Pintrich, 1997) und Lernstrategien, da diese für die vorliegende Arbeit von besonderem Interesse sind. Dazu wird im Rahmen dieses Kapitels knapp dargestellt, was Lernstrategien sind und welche Lernstrategieklassifikation dieser Arbeit zugrunde liegt, bevor empirische Ergebnisse zur Beziehung zwischen epistemologischen Überzeugungen und Lernstrategien beleuchtet werden.

Epistemologische Überzeugungen nehmen auf verschiedene Weise Einfluss auf den Lernprozess von Individuen. Sie beeinflussen alle Phasen des selbstregulierten Lernens und die metakognitiven Prozesse (z.B. Bromme, 2010; Bromme, Pieschl & Stahl, 2014; Muis, 2007; Muis & Edwards, 2009; Stahl, Pieschl & Bromme, 2006; Stallmann, 2007). Sie beeinflussen das Ausmaß, in dem Individuen aktiv lernen, bei schwierigen Aufgaben durchhalten und mit schlecht strukturierten Bereichen umgehen (vgl. Schommer, 1994). Sie wirken sich bei Lernenden aus auf Lernstrategieauswahl und -einsatz (z.B. Bråten & Strømsø, 2006; Garrett-Ingram, 1997; Kardash & Noel, 2000; M. P. Ryan, 1984; Schommer et al., 1992), die Textverarbeitung und das Textverstehen (z.B. Bråten & Strømsø, 2004; Kardash & Noel, 2000; M. P. Ryan, 1984; Schommer, 1990, 1994; Schommer et al., 1992; Schommer & Walker, 1995; Schommer-Aikins, 2004; Schraw et al., 2002; Strømsø & Bråten, 2011a) sowie auf die kognitive Verarbeitung (z.B. Kardash & Noel, 2000), die Motivation (z.B. Buehl & Alexander, 2004, 2005; Hofer, 1999; Muis & Foy, 2007; Paulsen & Feldman, 1999), die akademische Leistung (z.B. Cano, 2005; Garrett-Ingram, 1997; Hofer, 2000; Phillips, 2001; Schommer, 1993a; Schommer, Calvert, Gariglietti & Bajaj, 1997; Schommer-Aikins, 2004; Schommer-Aikins et al., 2005; Wood & Kardash, 2002) und den *Conceptual Change* (z.B. Andre & Windschitl, 2003; Mason, 2003; Mason & Boscolo, 2004; Qian & Alvermann, 1995; Qian & Pan, 2002; Windschitl & Andre, 1998). Auch bezüglich *moral reasoning* (z.B. Bendixen et al., 1998) und den Lernemotionen

(z.B. Muis et al., 2015) konnten Beziehungen zu den epistemologischen Überzeugungen gefunden werden.

Wie die Forschungsbefunde zeigen, können sich die individuellen epistemologischen Überzeugungen bezüglich all dieser lernrelevanten Aspekte entweder als förderlich oder aber hinderlich für das aktive Lernen erweisen (Schommer, 1994). Das bestehende wissensbezogene Überzeugungssystem eines Lernenden, das aus den spezifischen Ausprägungen in den verschiedenen Überzeugungsdimensionen besteht, kann das Lernen entweder unterstützen oder untergraben (z.B. Jonassen, Marra & Palmer, 2004; Moore, 1994; Palmer & Marra, 2008; Valanides & Angeli, 2005; Wineburg, 1991). Fortgeschrittenere, sophistiziertere Überzeugungen darüber, dass Wissen komplex und veränderlich ist, der Begründung und des Abwägens bedarf und in Interaktion konstruiert werden kann, gelten dabei größtenteils als lernförderlicher, während ein weniger elaboriertes epistemologisches Überzeugungssystem – vor allem bei komplexen Lernaufgaben – eher lernhinderlich ist (vgl. Bromme, 2010; Muis, 2004; Schommer, 1998; Schraw & Sinatra, 2004). Die epistemologischen Überzeugungen von Individuen entwickeln sich im Verlauf der akademischen Ausbildung und mit zunehmenden akademischen Erfahrungen weiter und werden sophistizierter (vgl. Bendixen & Rule, 2004; Muis et al., 2006 und siehe Kapitel 2.3.1). So gelten sophistizierte Überzeugungen gleichzeitig als wünschenswertes Ergebnis des Lehrens als auch als Prädiktoren für förderliches Lernverhalten und erfolgreiches Lernen (vgl. Hofer & Pintrich, 1997; Schoenfeld, 1992 sowie Trautwein et al., 2004). Die Effekte zwischen epistemologischen Überzeugungen scheinen reziprok zu sein. Sophistizierte epistemologische Überzeugungen führen zu förderlichem Lernverhalten (z.B. Conley et al., 2004; Gill et al., 2004; Hofer, 2004b; Mason & Boscolo, 2004), aber das resultierende lernförderliche Verhalten (z.B. der Einsatz von zur Lernaufgabe passenderen Lernstrategien) scheint auch mit einer Veränderung der epistemologischen Überzeugungen einherzugehen, die sich dann wiederum auf zukünftige Lerngelegenheiten auswirken (Schraw & Sinatra, 2004, S. 97). Es zeigt sich außerdem, dass diese Verbindung zwischen epistemologischen Überzeugungen und Lernen für die Lernenden verschiedenster Populationen und Entwicklungsstufen gilt (Schraw & Sinatra, 2004).

Sophistizierte epistemologische Überzeugungen unterstützen den Lerntransfer (Jacobson & Spiro, 1995) *Conceptual Change* (Franco et al., 2012; Mason & Boscolo, 2004; Qian & Alvermann, 1995), das Metaverständnis (Muis, 2007; Schommer et al., 1992) und den Umgang mit kontroversen Themen (Mason, 2000). Studierende mit sophistizierteren epistemologischen Überzeugungen können besser mit Fehlkonzepten umgehen und können diese eher ändern, als Studierende mit den gleichen Fehlvorstellungen aber weniger fortgeschrittenen Überzeugungen (Gallagher, 2002; Qian & Pan, 2002). Auch scheinen fortgeschrittenere epistemologische Überzeugungen im Zusammenhang mit Lernemotionen (z.B. Muis et al., 2015), intrinsischer Motivation (z.B. Buehl & Alexander, 2005; Hofer, 1999), Selbstkonzept (z.B. Urhahne & Harms, 2006), Selbststeuerungsfähigkeit (z.B. Bromme, 2010), Kontroll- und Selbstwirksamkeitsüberzeugung (z.B. Muis & Foy, 2007; Paulsen & Feldman, 1999) förderlicher zu sein.

In zwei Studien mit einer Stichprobe von insgesamt 439 deutschen, US-amerikanischen und kanadischen Studierenden erforschten Muis et al. (2015) die Verbindungen zwischen

epistemologischen Überzeugungen (TSEBQ, Bråten & Strømsø, 2010) und epistemischen Emotionen[6] sowie Lernstrategien. Sophistizierte epistemologische Überzeugungen wirkten sich förderlich auf die mit dem Lernen verbundenen Emotionen aus. Studierende, die davon überzeugt waren, dass die Rechtfertigung von Wissen zum Klimawandel das kritische Abwägen mehrerer Wissensquellen erfordert, erlebten stärker ausgeprägte Freude und Neugier und geringere Langeweile, wenn sie mit konfligierenden Informationen konfrontiert waren. Die sophistizierte Überzeugung, dass Wissen über Klimawandel komplex ist, ging mit geringerer Verwirrung, Angst und Langeweile einher. Die Überzeugung, dass Wissen unsicher ist, prädizierte niedrigere Angst und Frustration, während der Glaube an eine aktive Wissenskonstruktion niedrigere Verwirrung prädizierte. Es stellte sich heraus, dass die epistemischen Emotionen nicht nur die Arten der von den Studierenden eingesetzten Lernstrategien prädizierten, sondern den Zusammenhang zwischen epistemologischen Überzeugungen und Lernstrategien mediierten. Die Lernstratgien sagten wiederum den Lernerfolg vorher und mediierten die Verbindung zwischen epistemischen Emotionen und Lernerfolg (Muis et al., 2015).

Weniger fortgeschrittene Überzeugungen über die Einfachheit des Wissens und die angeborene Lernfähigkeit gingen bei Paulsen und Feldman (1999) mit einer geringen intrinsischen Zielorientierung, geringen Kontroll- und Selbstwirksamkeitsüberzeugungen sowie einer Geringschätzung von Lernaufgaben einher. Die Überzeugungen, dass Wissen einfach ist, ging außerdem mit höherer extrinsischer Zielorientierung und einer stärkeren Prüfungsangst einher. Köller et al. (2000) fanden positive Einflüsse der sophistizierten Überzeugungen auf das mathematische und physikalische Wissen von Schülerinnen und Schülern sowie auf deren fachspezifisches Selbstkonzept, Fachinteresse und Leistung. Qian und Pan (2002) kamen zu vergleichbaren Ergebnissen. Bei chinesischen und amerikanischen Studierenden korrelierte die weniger sophistizierte Überzeugung, dass Wissen einfach und sicher sowie die Lernfähigkeit angeboren ist, negativ mit der intrinschen Zielorientierung. Damit übereinstimmend fanden Bråten und Strømsø (2004) einen positiven Zusammenhang zwischen weniger sophistizierten Überzeugungen zur Sicherheit und Stabilität des Wissens sowie schnellem Lernen mit der Leistungszielorientierung und einen negativen Zusammenhang mit der Lernzielorientierung. Bei Neber und Schommer-Aikins (2002) korrelierte die Überzeugung von Hochbegabten an die angeborene Lernfähigkeit negativ mit der Aufgabenorientierung, während die Überzeugung zur Einfachheit des Wissens mit einer niedrigen Selbstwirksamkeit einherging. Die weiteren Zusammenhänge zwischen den epistemologischen Überzeugungen mit der Zielorientierung, Selbstwirksamkeit, intrinsischer Motivation und Prüfungsangst waren jedoch nicht signifikant. Buehl und Alexander (2004) untersuchten die Zusammenhänge zwischen den Überzeugungen zum mathematischen und geschichtlichen Wissen mit den Dimensionen des Erwartungs-Wert-Modells von (Wigfield & Eccles, 2000). Hier zeigen die Ergebnisse ein gemischtes Bild. Die wenig fortgeschrittene Überzeugung zur Sicherheit des Wissens korrelierte zwar positiv mit Kompetenz- und

[6]Epistemische Emotionen sind diejenigen, die im Individuum während des Lernens als Resultat der kognitiven Qualitäten von Aufgabeninformationen und der Verarbeitung dieser Informationen entstehen Pekrun, Goetz, Titz und Perry (2002a, 2002b) (*Epistemic Emotions Scales*, EES, Pekrun, Goetz, Frenzel, Barchfeld & Perry, 2011.

Valenzeinschätzungen, allerdings gingen die beiden ebenfalls wenig sophistizierten Überzeugungen zur Einfachheit und Quelle des Wissens mit niedrigen Kompetenzerwartungen und Valenzeinschätzungen für beide Fächer einher.

Epistemologische Überzeugungen von Lernenden haben direkte und indirekte Effekte auf die akademische Leistung. Indirekt wirken sie sich auf die Auswahl und Nutzung von Lernstrategien aus, während epistemologische Überzeugungen direkt als Filter für die Wahrnehmung der Lernumgebung wirken und so den Lernenden dazu veranlassen, beispielsweise Texte je nach eigener wissensbezogener Überzeugung anders zu interpretieren (Schommer, 1990, 1994; Schommer-Aikins & Hutter, 2002). Es gibt einige Befunde, die zeigen, dass sich sophistiziertere epistemologische Überzeugungen positiv auf die akademische Leistung von Studierenden auswirken (z.B. Hofer, 2000; Hofer & Pintrich, 1997; Schommer, 1990, 1993a, 1994; Schommer-Aikins & Hutter, 2002), bei Studierenden verschiedener Fachrichtungen mit einem besseren Notendurchschnitt assoziiert sind (Barnard et al., 2008) sowie mit höherer Motivation (Buehl & Alexander, 2005), im Sinne von stärker ausgeprägter Leistungszielorientierung (z.B. Hofer, 1999), Selbstwirksamkeit (z.B. Muis & Foy, 2007; Paulsen & Feldman, 1999) und fachspezifischem Selbstkonzept einhergehen (z.B. Urhahne & Harms, 2006). Auch bei Schülerinnen und Schülern der Sekundarstufe II und des Gymnasiums stellten sich sophistiziertere Überzeugungen (*Relativismus*) gegenüber weniger sophistizierten Überzeugungen (*Realismus* bzw. *Dualismus*) als vorteilhaft für die Leistung heraus (Baumert, Bos & Lehmann, 2000; Köller et al., 2000; Trautwein et al., 2004). Sophistiziertere Überzeugungen bei Schülerinnen und Schülern der Sekundarstufe I gingen bei Urhahne, Kremer und Mayer (2008) mit besseren kumulativen Lernleistungen in den Unterrichtsfächern Biologie, Chemie, Physik und Natur und Technik sowie mit einem höheren fachspezifischen Selbstkonzept in Physik bzw. Biologie einher. Schommer (1993a) untersuchte den Einfluss epistemologischer Überzeugungen auf die akademische Leistung an einer Stichprobe von 1182 amerikanischen Schülerinnen und Schülern einer High School (vom ersten Jahr bis zum vierten Jahr) (541 männlich, 641 weiblich). Es stellte sich einerseits heraus, dass die Überzeugungen zu den Dimensionen *simple knowledge*, *certain knowledge*, *quick learning* vom ersten bis zum vierten Jahr signifikant sophistizierter wurden. Mädchen neigten außerdem weniger als Jungen zu der Überzeugung, dass die Fähigkeit zu lernen angeboren ist und Lernen schnell erfolgt oder aber gar nicht. Die Überzeugung zu *quick learning* stellten sich, unter Berücksichtigung der allgemeinen Intelligenz, regressionsanalytisch als Prädiktor für die durchschnittliche akademische Leistung heraus. Je weniger die Schülerinnen und Schüler der Überzeugung waren, dass Lernen schnell oder aber gar nicht vonstatten geht, desto bessere Leistungen erreichten sie. Im Rahmen der TOSCA-Studie (Transformation des Sekundarschulsystems und akademische Karrieren, Trautwein & Lüdtke, 2004) wurde der Einfluss epistemologischer Überzeugungen auf den Lernerfolg mittels einer Stichprobe von Abiturientinnen und Abiturienten allgemein bildender Gymnasien untersucht. Sie setzten eine disziplinübergreifende Skala mit 15 Items zu Überzeugungen zur *Struktur des Wissens* ein. Faktorenanalytisch ergaben sich daraus die Subskalen *Dualismus* und *Relativismus*. In

Pfadanalysen konnte der Einfluss epistemologischer Überzeugungen auf die Abiturdurchschnittsnote bestätigt werden. Eine stärker dualistisch ausgeprägte Überzeugung zu Wissen ging mit schlechteren, hohe Relativismuswerte mit besseren Noten einher. Die epistemologischen Überzeugungen in dieser Stichprobe wurden durch Geschlecht, die kognitive Grundfähigkeit und Bildungsnähe der Familie moderiert. Die befragten Schülerinnen verfügten über niedrigere Dualismuswerte als die Schüler; und Schülerinnen und Schüler mit höherer kognitiver Grundfähigkeit und höherer Bildungsnähe wiesen niedrigere dualistische und höhere relativistische Überzeugungen auf. Die Überzeugung zur Sicherheit des Wissen beeinflusst nach den Ergebnissen von Kardash und Scholes (1996), wie Lernende Evidenzen zu kontroveresen Sachverhalten interpretieren. Tsai (1999) erforschte, wie die Überzeugungen Lernender ihre Beteiligung an spezifischen Lernaufgaben oder in Unterrichtssituationen beeinflussen und die Auswirkungen auf die Lernleistung. Eine konstruktivistische Sichtweise auf naturwissenschaftliches Wissen führte in ihrer Studie dazu, dass sich die entsprechenden Achtklässler eher darauf konzentrierten, die Bedeutung der durchgeführten Experimente untereinander auszuhandeln. Die Schülerinnen und Schüler mit empirischer Sichtweise auf Wissen setzten sich mit den Konzepten tiefer auseinander und erzielten ein besseres Verständnis.

Epistemologische Überzeugungen wirken sich, nach den Forschungsergebnissen von Stallmann (2007), Stahl et al. (2006), Bromme et al. (2014) und Bromme (2010), auf alle Phasen des selbstregulierten Lernens von Lernenden (Schülerinnen und Schülern, Studierenden) aus, indem sie die internen Standards für die metakognitive Überwachung und Regulation setzen und folglich den metakognitiven Anpassungsprozess an eine Lernaufgabe beeinflussen (Bromme, 2010). Die epistemologischen Überzeugungen beeinflussen die Standards, die sich ein Lernender bezüglich seiner Lernziele setzt. Aus den epistemologischen Überzeugungen ergeben sich spezifische Standards für das Wissen und den Wissensprozess, die wiederum einen Input für die metakognitiven Prozesse darstellen (Bromme, 2010; Muis, 2007). Epistemologische Überzeugungen wirken als individuelle (teilweise unterbewusste) Linse, durch die die Lernenden die Lernaufgabe wahrnehmen, die Lerninhalte erfassen (Bromme, 2010) und dienen der „calibration of learning to different learning tasks“ (Bromme, 2010, S. 7), d.h. der Kalibrierung von Lernstrategien unter Berücksichtigung der Anforderungen und Rahmenbedingungen, wie z.B. der Schwierigkeit und Komplexität der Aufgaben (Pieschl et al., 2013). Für einen Lernenden sind die eigenen Überzeugungen zur Beschaffenheit des Wissens untrennbar damit verbunden, wie er/sie eine Lernaufgabe eingeschätzt. Von den Überzeugungen zur Beschaffenheit des Wissens hängt ab, ob die Lernaufgabe beispielsweise als einfach oder komplex eingeschätzt wird; und daraus ergibt sich wiederum, wie der/die Lernende das eigene Lernen an die wahrgenommenen Aufgabeneigenschaften anpasst (Bromme, 2010). Es hat sich gezeigt, dass sophistiziertere epistemologische Überzeugungen mit mehr metakognitiven Überwachungsstrategien assoziiert sind (Bromme, 2010).

Epistemologische Überzeugungen beeinflussen das Verhalten von Lernenden bei der selbstregulierten Informationssuche und -bewertung (z.B. Hofer, 2004a; Mason et al., 2010b; Pieschl et al., 2013; Whitmire, 2004). Je fortgeschrittener die individuellen epistemologischen

Überzeugungen sind, desto eher wenden Studierende viel Zeit und Sorgfalt bei der Planung der Informationssuche auf, nutzen multiple Quellen, differenzieren zwischen der Herkunft von Inhalten und vergleichen kritisch zwischen verschiedenen Inhalten. Es hat sich gezeigt, dass die wissensbezogenen Überzeugungen filtern, wie Lehramtsstudierende das Lernen in ihren universitären Veranstaltungen erleben und inwiefern sie sinnvolle, tiefenorientierte Lernansätze betreiben (z.B. Many et al., 2002; Muis, 2004; Peng & Fitzgerald, 2006; Yadav & Koehler, 2007, vgl. Brownlee et al., 2011). Sie wirken folglich darauf, wie tief ihr in den universitären Veranstaltungen (*teacher education courses*) erworbenes Verständnis geht sowie auf ihre Entscheidungsfindung, Planung, Durchführung und Bewertung der praktischen Anwendung des Gelernten (Buehl & Fives, 2016; Yadav & Koehler, 2007).

Auch in Bezug auf das Arbeiten mit Texten, genauer die Textverarbeitung und das Textverstehen, hat sich mehrfach gezeigt, dass sophistizierte epistemologische Überzeugungen förderlich sind (Bråten & Strømsø, 2004; Kardash & Noel, 2000; Schommer et al., 1992; Schommer & Walker, 1995; Schommer-Aikins, 2004; Schraw et al., 2002). Schommer (1990) untersuchte die Effekte von epistemologischen Überzeugungen von College Studierenden ($N$ = 86) auf das Schlussfolgern, die Leistung in einem Wissenstest und die Überwachung des eigenen Verstehens beim Lesen von Texten (Themen Psychologie und Ernährung). Beim Lesen von Texten neigten Lernende mit der Überzeugung, dass Lernen schnell oder gar nicht vonstatten geht, dazu, stark vereinfachte Schlussfolgerungen aus den Texten zu ziehen. Außerdem erzielten sie eine geringere Textverstehensleistung, als die Lernenden mit sophistizierteren Überzeugungen, bei gleichzeitiger Überschätzung des eigenen Verstehens. Die Überzeugung, dass Wissen einfach ist, ging mit inadäquat absoluten Schlussfolgerungen zum Text einher. Übereinstimmend stellten Schommer et al. (1992) fest, dass Studierende, die der Überzeugung sind, dass Wissen objektiv ist, sich auf die Fakten des Textes konzentrieren und eingeschränkte Generalisierung und Transferanstrengung zeigen. Je weniger die Lernenden von der Einfachheit des Wissens überzeugt waren, desto bessere Verstehensleistungen erzielten sie und desto realistischer schätzten sie das eigene Verstehen ein.

Inwiefern ein weniger elaboriertes Überzeugungssystem bedeutende lernhemmende Einflüsse hat, zeigte Schommer (1998). Lernende, die der Überzeugung sind, dass Lernen schnell vonstatten geht, neigen demnach eher dazu, Texte oberflächlich zu lesen. Lernende, die glauben, dass Wissen sicher ist, neigen dazu, eher Fakten auswendig zu lernen als die Lerninhalte wirklich zu verstehen. Die Überzeugung, dass die Lernfähigkeit angeboren ist, geht mit weniger Interesse an Lernaktivitäten einher, die auf das Bewältigen komplexer Herausforderungen ausgerichtet sind. Während Lernende, die auf Autoritäten als Wissensübermittler vertrauen, dazu neigen, Wissensquellen nicht zu hinterfragen. Mason und Boscolo (2004) erforschten den Einfluss der epistemologischen Überzeugungen und des Themeninteresses von italienischen Schülerinnen und Schülern der zehnten und elften Jahrgangsstufe (38 weiblich, 27 männlich) auf die kritische Interpretation von Texten, in denen zwei verschiedene Positionen zu einem kontroversen Thema dargestellt sind. Außerdem interessierten sie sich dafür, ob das konzeptuelle Verstehen von den epistemologischen Überzeugungen beeinflusst wird und ob sich die Einstellung zu den Themen verändert und ob etwaige Veränderungen durch das epistemologische Verstehen und/oder das Interesse verursacht

werden. Sie stellten fest, dass die kritische Interpretation des im Text dargestellten kontroversen Themas nur von den epistemologischen Überzeugungen beeinflusst wird. Die Überzeugung, dass verschiedene Wissensurteile zu demselben Thema legitim sind, schien sich für die entsprechenden Lerner förderlich auf die kognitiven Prozesse des Identifizierens, Vergleichens, Schlussfolgerns und Bewertens der konkurrierenden Sichtweisen im Text auszuwirken. Schülerinnen und Schüler mit diesen fortgeschritteneren Überzeugungen erzielten im Vergleich zu solchen mit weniger fortgeschrittenen Überzeugungen bessere Bewertungen für ihre Antworten auf die offenen Fragen. Auch die Gruppe von Lernenden mit hohem Interesse am Thema erzielte bessere Leistungen als die der gering Interessierten. Bråten und Strømsø (2006) stellte fest, dass Lehramtsstudierende ($N$ = 39) mit sophistizierten epistemologischen Überzeugungen ein besseres Textverstehen beim Lesen multipler, teilweise konfligierender Texte erreichten, während die Lehramtsstudierenden mit weniger fortgeschrittenen Überzeugungen beim Lesen von einzelnen Texten ohne konfligierende Informationen ein besseres Verständnis erzielten. Bråten et al. (2008) konnten ebenfalls Verbindungen zwischen den epistemologischen Überzeugungen von Lehramtsstudierenden ($N$ = 135) und ihrem Textverstehen feststellen. Die Überzeugung, dass das Wissen zum Klimawandel komplex ist, führte zu besserem Textverstehen als die Überzeugung, dass Wissen zum Klimawandel einfach ist. Studierende, die in Autoritäten als Wissensvermittler zum Klimawandel vertrauten, erreichten allerdings interessanterweise bessere Ergebnisse als Studierende, die an eine Wissenskonstruktion glaubten. Franco et al. (2012) untersuchten die Rolle von epistemologischen Überzeugungen von Studierenden ($N$ = 75, 56 weiblich, 19 männlich) und der im Text vorherrschenden Wissensrepräsentation bezüglich der kognitiven und metakognitiven Verarbeitung beim Lernen mit Texten. Zur Erfassung der epistemologischen Überzeugungen diente der PEP von Royce und Mos (1980). Studierende, bei denen das Wissensrepräsentationsformat (rational vs. metaphorisch) des Textes konsistent zur individuellen epistemologischen Überzeugung (Empirismus, Metaphorismus) war, nutzen signifikant mehr tiefenorientierte kognitive Verarbeitungsstrategien beim Lesen der Texte und erzielten bessere Erinnerungsleistungen als Studierende mit zu den eigenen epistemologischen Überzeugungen inkongruenten Texten. Eine größere Anzahl von Studierenden der Gruppe mit zu den eigenen Überzeugungen kongruenten Texten konnte erfolgreich inhaltliche Misskonzepte abbauen (*Conceptual Change*). Die Studierenden erzielten bessere Lernleistungen, wenn die individuellen, wissensbezogenen Überzeugungen mit der Wissensrepräsentation im Physiktext übereinstimmten.

**Epistemologische Überzeugungen und Lernstrategien**

Unter Lernstrategien werden nach Weinstein und Mayer (1986, S. 315) die Verhaltensweisen und Kognitionen verstanden, die Lernende während des Lernens intentional zur Beeinflussung ihres Enkodierungsprozesses einsetzen. Nach diesem Lernbegriff erfordert der Lernprozess von einer lernenden Person die Selektion von Reizen im Sinne der aktiven Aufmerksamkeitszuwendung, die dauerhafte Speicherung von Informationen im Langzeitgedächnis, die Konstruktion von Verbindungen zwischen Konzepten und die Integration neuer Wissenselemente in die bereits bestehenden (Weinstein & Mayer, 1986). Lernstrategien

werden von Lernenden eingesetzt, um diese Anforderungen des Lernprozesses zu bewältigen und das Lernergebnis zu optimieren (ebd.). „Bei Lernstrategien handelt es sich (a) um eine Abfolge von effizienten Lerntechniken, die (b) zielführend und flexibel eingesetzt werden, (c) zunehmend automatisiert ablaufen, aber (d) bewusstseinsfähig bleiben." (Streblow & Schiefele, 2006, S. 353). Der Begriff Lerntechniken bezieht sind dabei auf konkrete Methoden, die beim Lernen eingesetzt werden, wie z.B. das Unterstreichen wichtiger Textstellen beim Lesen eines Textes. Setzt eine Person verschiedene Lerntechniken gezielt und koordiniert zur Unterstützung des Lernprozesses ein, spricht man von Lernstrategien (Streblow & Schiefele, 2006, S. 353). Lernstrategien stellen demnach Mediatoren zwischen Lernen und dem Lernergebnis dar (Biggs, 1993). Sie helfen dem Lernenden demnach bei der Bewältigung spezifischer Lernaufgaben (Wild, 2005, S. 193). Lernstrategien dienen sowohl der Steuerung kognitiver Prozesse als auch der Steuerung des Lernens (ebd.). Die Steuerung der kognitiven Prozesse beinhaltet Vorgehensweisen beim Wissenserwerb, die sich auf die Auswahl, den Erwerb, die Organisation und Integration von Informationen in bestehende Wissensstrukturen beziehen (Wild, 2000). Die Steuerung des Lernens erfolgt, indem die lernende Person die eigenen motivationalen und affektiven Zustände intentional beeinflusst (Wild, 2005, S. 193).

Welche Lernstrategien konkret unterschieden werden, ist von der zugrundeliegenden Konzeption abhängig. Überblicksarbeiten zur Klassifikation von Lernstrategien bieten Boekaerts, Pintrich und Zeidner (2000), Friedrich und Mandl (1992), Wild und Schiefele (1993) sowie Wild (2000). Eine bedeutende Lernzielklassifikation, die eine Differenzierung und Weiterentwicklung der Konzeption von Weinstein und Mayer (1986) darstellt, stammt von der Forschergruppe um Pintrich und McKeachie (siehe Wild, 2005, S. 193). Aus dieser ging der *Motivated Strategies for Learning Questionnaire* (MSLQ) von Pintrich et al. (1991) hervor. In Abbildung 2.10 auf Seite 98 ist diese Lernzielklassifikation von Pintrich et al. (1991) dargestellt. Wie hier zu sehen ist, gliedert sich der Bereich der Lernstrategien in die drei übergeordneten Bereiche *kognitive Lernstrategien*, *metakognitive Lernstrategien* und *Strategien zur Nutzung interner und externer Ressourcen* mit insgesamt 13 Lernstrategiedimensionen. Die kognitiven Lernstrategien werden als Primärstrategien bezeichnet. Sie dienen direkt der Informationsaufnahme, -verarbeitung und -speicherung (Wild, 2005, S. 194). Darunter fallen *Wiederholungs-*, *Organisations-* und *Elaborationsstrategien* sowie *kritisches Denken*. Die Wiederholungsstrategien zielen darauf ab, Lerninhalte auswendig zu lernen, indem Fakten anhand von Listen oder Zusammenfassungen immer wieder (laut oder leise) aufgesagt werden. Im Rahmen der Elaboration von Lerninhalten kann ein Lernender neue Informationen mit dem Vorwissen in Beziehung setzen, sie in eigene Worte fassen, durch Beispiele illustrieren und kritisch prüfen (Wild, 2000). Die Organisationsstrategien umfassen Methoden, mittels derer Kernaussagen aus Texten herausgearbeitet werden und Inhalte ausgewählt und neu strukturiert werden, wie z.B. durch die Anfertigung von Tabellen und Mindmaps (Streblow & Schiefele, 2006, S. 354). Beim kritischen Denken nutzen Lernende ihr Vorwissen, indem sie es „auf neue Situationen anwenden, um Probleme zu lösen, Entscheidungen zu treffen oder kritische Bewertungen vorzunehmen" (Wild, 2000, S. 38).

| | | | |
|---|---|---|---|
| Primärstrategien | Kognitive Lernstrategien | oberflächenorientiert | Wiederholungsstrategien |
| | | tiefenorientiert | Organisationsstrategien |
| | | | Elaborationsstrategien |
| | | | Kritisches Denken |
| | Metakognitive Lernstrategien | | Lernschritte planen |
| | | | Lernschritte überwachen |
| | | | Lernschritte regulieren |
| Sekundärstrategien | Strategien zur Nutzung interner und externer Ressourcen | interne | Anstrengung |
| | | | Aufmerksamkeit |
| | | | Zeitmanagement |
| | | externe | Gestaltung von Lernumgebung und Arbeitsplatz |
| | | | Nutzung von Informationsquellen |
| | | | Lernen mit Studienkollegen |

ABBILDUNG 2.10: Darstellung der Lernstrategieklassifikation des MSLQ von Pintrich et al. (1991, S. 194), in eigener Darstellung

Wiederholungsstrategien stellen eine oberflächenorientierte kognitive Strategieart dar, da sie auf eine oberflächliche kognitive Verarbeitung im Sinne des reinen Behaltens aber nicht Verstehens von Lerninhalten abzielen. Die Organisations- und Elaborationsstrategien sowie das kritische Denken sind tiefenorientierte Strategien. Sie sind auf das Verstehen der Lerninhalte ausgerichtet. Neue Inhalte werden verarbeitet, indem sie umstrukturiert, auf Wesentliches reduziert und mit dem Vorwissen verknüpft werden (Streblow & Schiefele, 2006, S. 354). Die metakognitiven Strategien umfassen das *Planen*, *Überwachen* und *Regulieren* der Lernschritte. Die Strategien zur Nutzung interner und externer Ressourcen, die als Sekundärstrategien beschrieben werden, umfassen je drei externe und interne Ressourcen. Die externen Ressourcen sind *Gestaltung von Lernumgebung und Arbeitsplatz* (z.B. wichtige Unterlagen bereit legen) ), *Nutzung von Informationsquellen* (z.B. Nutzen von Literatur zum Nachschlagen von Inhalten) und *Lernen mit Studienkollegen* (z.B. Inhalte mit Kommilitonen durchsprechen). Die internen Ressourcen betreffen die *Anstrengung* (z.B. auch bei mangelndem Interesse die Anstrengung aufrechterhalten), *Aufmerksamkeit* (z.B. das Abschweifen der Gedanken verhindern) und das *Zeitmanagement* (z.B. sich beim Lernen an einen Zeitplan halten) (siehe Streblow & Schiefele, 2006, S. 354 sowie Wild, 2000). Diese Konzeption des MSLQ (siehe Abb. 2.10) liegt auch einigen Adaptionen dieses Fragebogens zugrunde. Unter anderem auch dem für den Hochschulbereich bestimmten *Inventar zur Erfassung von Lernstrategien im Studium* (LIST) von Wild und Schiefele (1994) sowie Wild (2000), das inzwischen in einer revidierten Fassung (LIST-R, Wild, 2011) vorliegt.

Es gibt zahlreiche Studien, die Zusammenhänge zwischen den epistemologischen Überzeugungen und der Lernstrategieauswahl bzw. teilweise auch dem Lernstrategieeinsatz

fanden sowie Effekte der epistemologischen Überzeugungen auf die Bevorzugung von bestimmten Lernstrategiearten zeigen. Häufig konnten bei Studierenden positive Zusammenhänge zwischen sophistizierten epistemologischen Überzeugungen und tiefenorientierten Lernstrategien gefunden werden. Also solchen Lernstrategien, die auf das Verstehen der Inhalte abzielen, statt auf das Auswendiglernen. Während die unsophistizierteren epistemologischen Überzeugungen eher positive Zusammenhänge mit den oberflächenorientierten Lernstrategien aufweisen.

Für den Lernprozess von Schülerinnen und Schülern im Unterricht wurde die Bedeutung epistemologischer Überzeugungen vielfach untersucht. Dweck und Legett (1988) stellten fest, dass Kinder, die der starken Überzeugung sind, dass die Fähigkeit zu lernen, veränderbar ist, Lernaufgaben als Herausforderung wahrnehmen, verschiedene Lernstrategien ausprobieren und über mehr Durchhaltevermögen beim Lernen verfügen, als Kinder, die an eine angeborene Lernfähigkeit glauben. Wenn zweitere mit schwierigen Aufgaben konfrontiert werden, zeigen sie hilfloses Verhalten. In einer qualitativen Studie mit Schülerinnen und Schülern stellte Tsai (1998) fest, dass fortgeschrittenere Überzeugungen (*empiricist beliefs*) zum naturwissenschaftlichen Wissen und Wissensprozess mit der Neigung zu Lernstrategien einhergehen, die auf das Verstehen abzielen. Weniger fortgeschrittene epistemologische Überzeugungen dagegen hingen mit Lernstrategien zusammen, die auf das Auswendiglernen abzielen.

Die Modellannahmen Hofers (2001) zu der Beziehung zwischen den individuellen epistemologischen Überzeugungen und verschiedenen lernrelevanten Konstrukten (u.a. Lernstrategien) werden durch die Ergebnisse von Urhahne und Hopf (2004) gestützt. Unter Einsatz einer deutschen Version des Fragebogens von Conley et al. (2004) bei 167 Schülerinnen und Schülern (102 männlich, 65 weiblich) der neunten Jahrgangsstufe ($M$ = 15.36, $SD$ = .61) dreier Gymnasien ermittelten Urhahne und Hopf (2004) Zusammenhänge zwischen den epistemologischen Überzeugungen (Quelle, Entwicklung, Sicherheit, Rechtfertigung des Wissens) und Leistungsmotivation, fachspezifisches Selbstkonzept in Biologie und Physik sowie dem Lernstrategieeinsatz. Die genutzen Lernstrategieskalen (Wiederholung, Elaboration, metakognitive Kontrolle, Anstrengung nach Wild & Schiefele, 1994) stammen aus der Dokumentation der Erhebungsinstrumente aus PISA 2000 (Kunter et al., 2003). Wie aus Tabelle 2.4 auf Seite 100 hervorgeht, gingen fortgeschrittenere Überzeugungen zur Dimensionen *Rechtfertigung des Wissens* erwartungskonform mit einem signifikant stärkeren Einsatz (Selbstauskunft) der anspruchsvolleren Elaborations- und Kontrollstrategien einher. Schülerinnen und Schüler mit fortgeschrittenen Überzeugungen zur Rechtfertigung des Wissens sind darüber hinaus leistungsmotivierter und verfügen über ein höheres naturwissenschaftliches Interesse sowie fachspezifisches Selbstkonzept in Physik und Biologie (siehe Tab. 2.4, S. 100). Schülerinnen und Schüler, die der sophistizierteren Überzeugung sind, dass Wissen unsicher ist, berichten, weniger Memorierstrategien einzusetzen. Fortgeschrittene Überzeugungen zur Entwicklung des Wissens korrelierten positiv mit dem berichteten Einsatz von metakognitiven Kontrollstrategien sowie mit Leistungsmotivation und dem fachspezifischen Selbstkonzept in Biologie (siehe Tab. 2.4, S. 100). Die fortgeschrittene Überzeugung bezüglich der Dimension Quelle des Wissens wiesen wider Erwarten einen negativen

Zusammenhang mit Elaborations-, Kontroll- und Anstrengungsstrategien sowie naturwissenschaftlichem Interesse auf (siehe Tab. 2.4). Das bedeutet bezüglich der Lernstrategien, dass Schülerinnen und Schüler, die Autoritäten als Quelle des Wissens sehen, eher bereit sind „sich anzustrengen, ihre Lernprozesse zu kontrollieren und zu versuchen, inhaltlich Neues mit bereits Bekanntem zu verbinden." (Urhahne & Hopf, 2004, S. 80).

TABELLE 2.4: Bivariate Korrelationen zwischen epistemologischen Überzeugungen, Motivation, Lernstrategien und Selbstkonzept nach Urhahne und Hopf (2004, S. 79)

| Variable | Rechtfertigung | Quelle | Entwicklung | Sicherheit |
|---|---|---|---|---|
| Wiederholungsstrategien | -.14 | -.03 | -.12 | -.19* |
| Elaborationsstrategien | .34*** | -.22** | .12 | -.11 |
| Kontrollstrategien | .35*** | -.26** | .21** | -.05 |
| Anstrengungsbereitschaft | .13 | -.23** | .05 | -.12 |
| Naturwiss. Interesse | .22** | -.25** | .10 | -.20* |
| Leistungsmotivation | .22** | -.04 | .27** | .10 |
| Selbstkonzept Biologie | .39*** | -.08 | .44*** | .15 |
| Selbstkonzept Physik | .28** | -.13 | .18* | .12 |

*$p < .05$, **$p < .01$, ***$p < .001$

Fortgeschrittenere epistemologische Überzeugungen korrelieren positiv mit Leistungsmotivation und den fachspezifischen Selbstkonzepten in Biologie und Physik (Urhahne & Hopf, 2004). Da Urhahne und Hopf (2004) nur bivariate Zusammenhänge zwischen epistemologischen Überzeugungen und den lernrelevanten Faktoren untersuchen, sind keine Aussagen über Wirkrichtungen zulässig.

Cano (2005) geht von direkten als auch indirekten Einflüssen epistemologischer Überzeugungen auf die Lernleistung aus. Der indirekte Einfluss wird dabei entsprechend der Modelle von Hofer (2001) und Schommer-Aikins (2004) von den Lernstrategien mediiert. Cano (2005) prüfte die angenommmen Beziehungen mittels Strukturgleichungsmodellen anhand einer Stichprobe von 1600 spanischen Schülerinnen und Schülern (722 weiblich, 878 männlich) im Alter zwischen zwölf und zwanzig ($M$ = 14.79, $SD$ = 1.95). Die epistemologischen Überzeugungen wurden mittels einer ins Spanisch übersetzten Version des SEQ (Schommer, 1990) erfasst, während zur Erfassung der Lernstrategien (*approaches to learning*) eine spanische Version des *Study Process Questionnaire* (SPQ) von Biggs (1987) eingesetzt wurde. Auf Basis der faktorenanalytischen Ergebnisse wurde zwischen tiefenorientierten und oberflächenorientierten Lernstrategien (*approaches to learning*) unterschieden. In einem Pfadmodell stellt Cano (2005, S. 214) die signifikanten Pfade als standardisierte Koeffizienten dar. Dieses Pfadmodell ist in Abbildung 2.11 auf Seite 101 zu sehen ($\chi^2$ = 5.19 (df = 3), $p$ = .16, GFI = 1.00, AGFI = .99 und SRMR = .01). Lernende, die tiefenorientierte Lernstrategien bevorzugen, versuchen die großere Bedeutung und Intention eines Textes zu verstehen und die Lerninhalte mit dem Vorwissen zu verknüpfen (siehe Abb. 2.11, S. 101). Lernende, die oberflächenorientierte Lernstrategien bevorzugen, zielen eher auf das Memorieren der Inhalte statt auf das Verstehen und achten demnach mehr auf Details und Fakten von

beispielsweise Texten. Erwartungskonform beeinflussten die epistemologischen Überzeugungen (*simple knowledge*, *quick learning*) einerseits direkt die akademische Leistung, die über die durchschnittliche Schuljahresabschlussnote operationalisiert wurde. Andererseits wirkten sie sich auch auf die bevorzugten Lernstrategien aus (siehe Abb. 2.11). Weniger fortgeschrittene epistemologische Überzeugungen bezüglich *simple knowledge* und *quick learning* führten zu einer schlechteren akademischen Leistung. Je mehr ein Lernender glaubte, dass Lernen schnell und ohne Anstrengung erfolgt, desto wahrscheinlicher neigte er/sie zu oberflächenorientierten Strategien und desto weniger zu Tiefenstrategien (Cano, 2005). Die naive Überzeugung, dass Wissen einfach ist, prädizierte sowohl Oberflächen- als auch Tiefenstrategien, allerdings erstere stärker. Die naive Überzeugung, dass Wissen sicher ist, prädizierte in geringem Ausmaß die Oberflächenstrategien. Die tiefenorientierten Lernstrategien beeinflussen die schulische Leistung positiv, während die oberflächenorientierten Lernstrategien diese negativ beeinflussten.

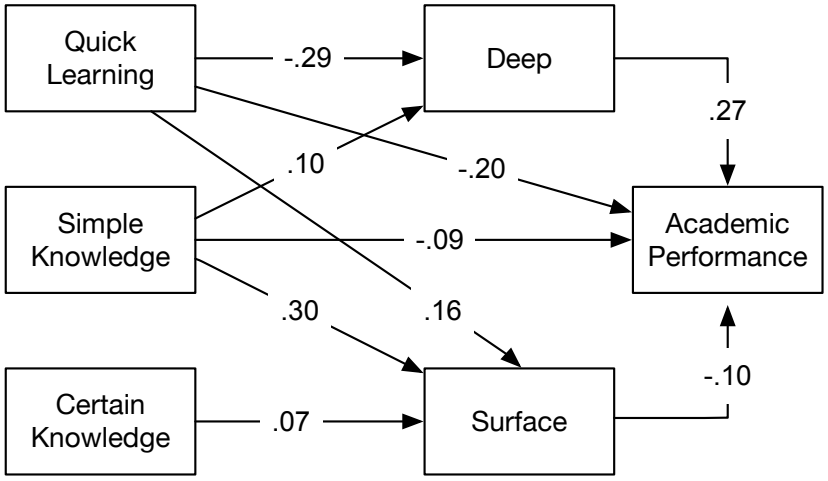


ABBILDUNG 2.11: Pfadmodell zum Einfluss epistemologischer Überzeugungen auf *approaches to learning* und die akademische Leistung (Cano, 2005, S. 214) in eigener Darstellung

Die Verbindung zwischen den epistemologischen Überzeugungen und dem Lernprozess von Studierenden ist ebenfalls gut untersucht und führt zu vergleichbaren Ergebnissen. So untersuchte beispielsweise M. P. Ryan (1984), inwieweit sich Studierende ($N$ = 90, 58 weiblich, 33 männlich) mit unterschiedlichen epistemologischen Standards im Hinblick auf die Anwendung der Verstehenskriterien der Bloomschen Taxonomie (Bloom, Engelhart, Furst, Hill & Krathwohl, 1956) beim Lesen von Texten unterscheiden. Die Studierenden wurden auf Basis von Perrys Einstellungsitems (Perry, 1968) in die Gruppe dualistisches (faktenorientiert) oder relativistisches (kontextorientiert) Wissenskonzept aufgeteilt. Er stellte fest, dass Studierende mit dualistisch ausgeprägten epistemologischen Überzeugungen beim Lesen der Texte die reine Prüfung des Textverständnisses (Wissenskriterium) bevorzugten, was der Anwendung der Wiederholungsstrategie entspricht. Studierende mit relativistischer Perspektive auf Wissen wendeten häufiger Kriterien des Verstehens und Anwendens an. Sie wendeten Strategien an, um nachvollziehbare Relationen zwischen Textaussagen

herzustellen. Studierende, die beim Lesen eher Kriterien des Verstehens und Anwendens anwandten, erzielten bessere Noten, als die Studierende, denen das Faktenwissen wichtiger war (M. P. Ryan, 1984, S. 252 f.).

Schommer et al. (1992) fanden Hinweise darauf, dass der Einfluss der individuellen epistemologischen Überzeugungen auf das Textverstehen und die adäquate Einschätzung des eigenen Verstehens durch die Lernstrategien mediiert wird. Richter (2003) untersuchte, welche epistemologischen Strategien von Studierenden im akademischen Lernen mit Texten eingesetzt werden ($N$ = 289 Studierende). Anwendung von Überzeugungswissen, Konsistenzprüfung, überlegte Bewertung, Leichtigkeit und Sicherheit der Textbewertung sowie wissensgestützte Bewertungen (epistemologische Einschätzungen zum Textverstehen) haben bedeutenden Einfluss darauf, wie Lernende Sachtexte verstehen und dass eine epistemologische Verarbeitung einer rezeptiven Verarbeitung überlegen ist. Richter geht davon aus, dass adäquates Textverständnis nur erreicht wird, wenn ein Text auch epistemologisch verarbeitet wird.

Urhahne und Harms (2006) übertrugen die Modellannahmen Hofers (2001) auf das akademische Lernen von Studierenden und untersuchten bei diesen die Bedeutung disziplinspezifischer epistemologischer Überzeugungen für die Lernstrategien sowie für Selbstkonzept und Motivation. Zur Erfassung der Überzeugungen zum naturwissenschaftlichen Wissen und Wissensprozess wurde der von Urhahne und Hopf (2004) ins Deutsche übersetzte Fragebogen von Conley et al. (2004) mit den Skalen *Entwicklung*, *Sicherheit*, *Quelle* und *Rechtfertigung des Wissens* eingesetzt. Die Stichprobe bestand aus 120 Lehramtsstudierenden (103 weiblich, 17 männlich) im durchschnittliche Alter von 23.12 ($SD$ = 5.09) des Fachs Biologie. Es wurden die Lernstrategieskalen Wiederholungs-, Elaborations-, metakognitive Kontrollstrategien nach Wild und Schiefele (1994) aus der Dokumentation der Erhebungsinstrumente aus PISA 2000 (Kunter et al., 2003) entnommen. Wie Tabelle 2.5 zu entnehmen, gingen allgemein weniger entwickelte epistemologische Überzeugungen erwartungskonform mit vermehrten Wiederholungsstrategien einher.

TABELLE 2.5: Bivariate Korrelationen zwischen epistemologischen Überzeugungen, Motivation, Lernstrategien und Selbstkonzept nach Urhahne und Harms (2006, S. 193)

| Variable | Sicherheit | Quelle | Rechtfertigung | Entwicklung |
|---|---|---|---|---|
| Wiederholungsstrategien | -.29** | -.39** | .01 | -.26** |
| Elaborationsstrategien | .08 | .05 | .24** | .22* |
| Kontrollstrategien | .04 | -.05 | .23* | .15 |
| | | | | |
| Naturwiss. Interesse | -.13 | -.08 | .24** | .13 |
| Leistungsmotivation | .13 | .14 | .22* | .20* |
| Selbstkonzept Biologie | -.03 | .02 | .33*** | .21* |

*$p < .05$, **$p < .01$, ***$p < .001$

Studierende, die stärker von der Sicherheit und Unveränderlichkeit des naturwissenschaftlichen Wissens überzeugt sind sowie von der Rolle der Autoritäten als Wissensvermittler, geben an, beim Lernen stärker Wiederholungsstrategien zu nutzen, als Studierende

mit fortgeschrittenen wissensbezogenen Überzeugungen (siehe Tab. 2.5, S. 102). Die fortgeschrittene Ansicht, dass naturwissenschaftliches Wissen sich weiterentwickelt, korreliert positiv mit der Angabe, Elaborationsstrategien beim Lernen einzusetzen. Studierende, die der fortgeschrittenen Überzeugung sind, dass naturwissenschaftliches Wissen zu begründen und zu rechtfertigen ist, favorisieren Elaborationsstrategien und metakognitive Kontrollstrategien und sind naturwissenschaftlich interessierter beim Lernen. Fortgeschrittenere epistemologische Überzeugungen (Entwicklung, Rechtfertigung) gingen außerdem mit höherer Leistungsmotivation und einem höheren fachspezifischen Selbstkonzept einher (siehe Tab. 2.5, S. 102).

In einer anschließenden Clusteranalyse konnten Urhahne und Harms (2006) vier verschiedene Überzeugungstypen identifiziert werden, zwischen denen signifikante Unterschiede bezüglich der Neigung zu Wiederholungsstrategien und Elaborationsstrategien sowie im fachspezifischen Selbstkonzept und naturwissenschaftlichen Interesse gefunden wurden. Personen mit einer relativistischen Weltsicht (sophistiziertere epistemologische Überzeugungen) verfügten über ein signifikant höheres Selbstkonzept und setzten nach eigenen Angaben signifikant häufiger Elaborationsstrategien und weniger Wiederholungsstrategien ein als Personen mit wenig fortgeschrittenen wissensbezogenen Überzeugungen (Sicherheit, Konstanz des Wissens und Autoritäten als Wissensvermittler).

Phan (2006) untersuchte mittels Strukturgleichungsmodellen die Beziehungen zwischen epistemologischen Überzeugungen, Lernstrategien, Lernleistung und reflektiertem Denken bei 332 Studierenden (146 weiblich, 185 männlich) der Pädagogik im durchschnittlichen Alter von 22 Jahren ($SD$ = 1.12). Zur Erfassung der epistemologischen Überzeugungen wurde der SEQ (Schommer, 1990) eingesetzt, während die Lernstrategien mittels des revidierten *2- Factors-Study-Progress-Questionnaire* (R-SPQ-2F) von Biggs, Kember und Leung (2001) erhoben wurden. Der *Reflective Thinking Questionnaire* (RTQ) von Kember et al. (2000) mit den vier Stufen reflektierten Denkens *habitual action*, *understanding*, *reflection* und *critical reflection* nach (Mezirow, 1991, 1998) diente zur Erfassung des reflektierten Denkens. Während *habitual action* eine einfache und fast automatisierte Informationsaufnahme bezeichnet, bei der das Gelernte nicht bewusst verarbeitet wird, umfasst *understanding* die Verarbeitung des Gelernten, allerdings ohne Verknüpfungen zu anderen Wissensinhalten herzustellen. Werden aktiv und sorgfältig alle Annahmen berücksichtigt, die dem Lerninhalt zugrunde liegen, spricht man von *reflection*. *Critical reflection* stellt eine weiterentwickelte Stufe von *reflection* dar. Diese Stufe des reflexiven Denkens umfasst ein Bewusstsein für die Ursachen von Wahrnehmungen, Gefühlen und Handlungen. Das resultierende Modell ist in Abbildung 2.12 auf Seite 104 dargestellt ($\chi^2$ = 39.64, df = 12, p < .01, CFI = .98, NNFI = .92, RMSEA = .08). Hier sind nur die signifikanten Pfade aufgeführt. Die Überzeugungen zu *simple knowledge* und *fixed ability* wirkten sich einerseits direkt auf die Leistung zum Ende des Semesters aus (siehe Abb. 2.12, S. 104). Überzeugungen zu *simple knowledge* beeinflussten andererseits sowohl die Wahl von tiefenorientierten und oberflächenorientierten Lernstrategien. Die tiefenorientierten Strategien wirkten sich wiederum auf die Lernleistung aus. Überzeugungen darüber, ob die Fähigkeit zu lernen angeboren ist oder aber erlernt werden kann, wirkte sich auch direkt auf die Reflexion und das kritische Denken aus. Für *quick learning* und *certain*

*knowledge* wurden in dieser Studie keine Effekte auf die Lernstrategien oder die Lernleistung festgestellt (siehe Abb. 2.12).

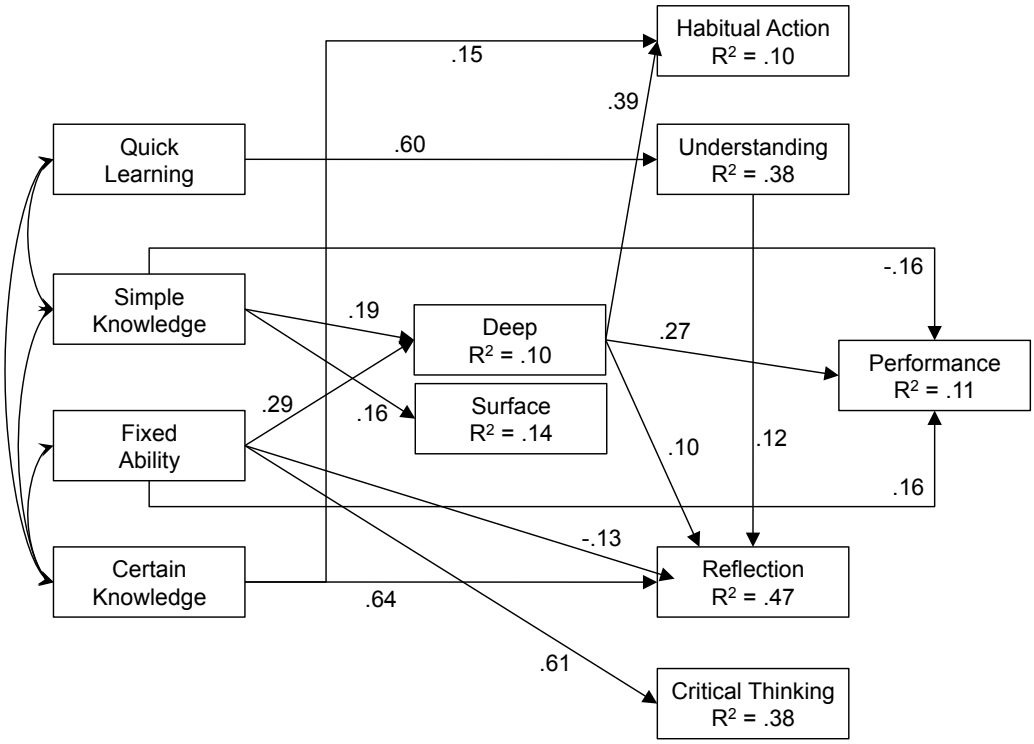


ABBILDUNG 2.12: Pfadmodell zum Einfluss epistemologischer Überzeugungen auf Lernstrategien, Verstehen, kritisches Denken, Reflexion, Habitual Action und Lernleistung von Phan (2006)

Phan (2008) untersuchte Veränderungen in epistemologischen Überzeugungen, Lernstrategien und reflektiertem Denken an zwei Stichproben von 616 Studierende im ersten Studienjahr Kunst (217 weiblich, 345 männlich) und 581 Mathematikstudierenden (241 weiblich, 341 männlich) im zweiten Studienjahr. Die epistemologischen Überzeugungen wurden mittels des SEQ (Schommer, 1990) und die Lernstrategien mittels des R-SPQ-2F von Biggs et al. (2001) erfasst. Zur Erfassung der Lernstrategien diente der R-SPQ-2F von Biggs et al. (2001) mit den Skalen *Deep Motive*, *Deep Strategy*, *Surface Motive* und *Surface Strategy*. Die Analyse erfolgte für beide Stichproben getrennt mittels Strukturgleichungsmodellen (*cross-lag model* und *causal-mediating model*). Sie ergab für die Kunststudierenden, dass die epistemologischen Überzeugungen zum ersten Messzeitpunkt (t1) die Lernstrategien zum zweiten Messzeitpunkt (t2) prädizieren ($\beta$ = .09), allerdings die Lernstrategien (t1) umgekehrt nicht die epistemologischen Überzeugungen (t2) ($\beta$ = .03, ns). Sowohl die berichteten Lernstrategien als auch die epistemologischen Überzeugungen stellten sich als relativ stabil über den Zeitraum von einem Jahr heraus. Die epistemologischen Überzeugungen zu t1 sagten die Überzeugungen zu t2 vorher ($\beta$ = .52) und die Strategien zu t1 die zu t2 ($\beta$

= .58). Anhand des Modells für die Stichprobe der Mathematikstudierenen konnte einerseits gezeigt werden, dass die epistemologischen Überzeugungen die Lernstrategien ($\beta$ = .53), das reflektierte Denken ($\beta$ = .35) und die akademische Leistung ($\beta$ = .19) prädizieren. Andererseits sagten auch die Lernstrategien die akademische Leistung ($\beta$ = .14) sowie das reflektierte Denken ($\beta$ = .38) vorher. Phan (2008) schlussfolgert aus den Ergebnissen, dass die „students' learning approaches are a product of their epistemological beliefs. Furthermore, the finding supports the contention that students' SAL (approaches to learning) mediates the effect of epistemological beliefs on academic performance and reflective thinking. " (S. 89).

Dahl et al. (2005) untersuchten regressionsanalytisch die Beziehung zwischen epistemologischen Überzeugungen und Lernstrategien mittels einer Stichprobe von 81 Studierenden (21 männlich, 60 weiblich) im Durchschnittsalter von 22.6 Jahren (*SD* = 6.4). Die epistemologischen Überzeugungen zu *Simple Knowledge*, *Certain Knowledge*, *Fixed Ability* und *Quick Learning* wurden mittels des SEQ (Schommer, 1990; Schommer & Walker, 1995) erfasst. Die Lernstrategien wurden über den MSLQ (Pintrich et al., 1991) erhoben (Skalen: *Wiederholungsstrategien*, *Elaborierungsstrategien*, *Strategien des kritischen Denkens*, *Organisationsstrategien* sowie *metakognitive und selbstregulative Strategien*). In Tabelle 2.6 sind zunächst die bivariaten Zusammenhänge zwischen den Skalen der epistemologischen Überzeugungen und den Lernstrategieskalen dargestellt.

TABELLE 2.6: Bivariate Korrelationen zwischen epistemologischen Überzeugungen und Lernstrategien (Dahl et al., 2005, S. 265)

| Skalen | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1. Simple | | | | | | | | |
| 2. Fixed | .29** | | | | | | | |
| 3. Quick | .18 | .27* | | | | | | |
| 4. Certain | .06 | .37** | .03 | | | | | |
| 5. Rehearsal | -.31** | -.16 | .03 | -.09 | | | | |
| 6. Elaboration | -.27* | -.44** | -.16 | -.13 | .41** | | | |
| 7. Critical Thinking | -.15 | -.50** | -.13 | -.17 | .26* | .59** | | |
| 8. Organization | -.43** | -.32** | -.04 | -.22* | .55** | .62** | .42** | |
| 9. Metacognitive Self-regulation | -.41** | -.35** | -.08 | -.15 | .56** | .62** | .53** | .66** |

*$p < .05$, **$p < .01$, *** $p < .001$

Hier zeigt sich, dass die Überzeugung in eine einfache Struktur des Wissens (*Simple Knowledge*) signifikant negativ mit Wiederholungs-, Elaborations-, Organisations- und den metakognitiven selbstregulativen Strategien korreliert (siehe Tab. 2.6). Die Überzeugung, Wissen sei sicher, geht mit weniger Organisationsstrategien einher. Die Überzeugung in eine angeborene Fähigkeit zu Lernen (*Fixed Ability*) geht mit den Überzeugungen einher, Lernen gehe schnell oder aber gar (*Quick Learning*) nicht sowie Wissen sei sicher (*Certain Knowledge*). *Fixed Ability* hängt signifikant negativ mit Elaborations-, Organisations- und den metakognitiven Strategien sowie kritischem Denken zusammen.

In den Tabellen 2.7 und 2.8 sind die Ergebnisse der Regressionsanalyse zu den epistemologischen Überzeugungen auf die Lernstrategien dargestellt. Nur die Prädiktoren *Fixed Ability* und *Simple Knowledge* stellten sich als signifikant heraus. Die Wiederholungsstrategien, Organisationsstrategien und metakognitiven und selbstregulativen Strategien werden durch *Simple Knowledge* prädiziert. *Fixed Ability* ist ein signifikanter Prädiktor für die Elaborationsstrategien und metakognitiven und selbstregulativen Strategien. Es werden allerdings über die verschiedenen Regressionsmodelle lediglich Varianzaufklärungen von 11% bis maximal 25% erreicht, d.h. die epistemologischen Überzeugungen können nur einen geringen Teil der Lernstrategieneigung erklären. Zu berücksichtigen ist allerdings hier die geringen Stichprobengröße, die zur ineffizienten Schätzung der Parameter führen kann.

TABELLE 2.7: Multiple Regressionen epistemologischer Überzeugungen auf Lernstrategien (Dahl et al., 2005, S. 267 ff.) – Teil 1

| | Wiederholung | | Elaboration | | Kritisches Denken | |
|---|---|---|---|---|---|---|
| Prädiktor | beta | t | beta | t | beta | t |
| Simple | -.30 | -2.60** | -.15 | -1.43 | .00 | .00 |
| Fixed | -.08 | -.64 | -.39 | -3.37*** | -.50 | -4.37*** |
| Quick | .10 | .89 | -.03 | -.27 | .00 | .04 |
| Certain | -.04 | -.36 | .03 | .28 | .02 | .16 |
| $R^2$ | .11 | | .22 | | .25 | |
| F | 2.33 | | 5.24*** | | 6.22*** | |

**$p < .01$, *** $p < .001$

TABELLE 2.8: Multiple Regressionen epistemologischer Überzeugungen auf Lernstrategien (Dahl et al., 2005, S. 267 ff.) – Teil 2

| | Organisation | | metakognitive/ selbstregulative | |
|---|---|---|---|---|
| Prädiktor | beta | t | beta | t |
| Simple | -.38 | -3.64** | -.34 | -3.21** |
| Fixed | -.18 | -1.52 | -.25 | -2.16 |
| Quick | .08 | .79 | .05 | .46 |
| Certain | -.14 | .12.5 | -.04 | -.33 |
| $R^2$ | .25 | | .23 | |
| F | 6.16*** | | 5.65*** | |

**$p < .01$, *** $p < .001$

Auf Basis des COPES-Modells zum selbstregulierten Lernen und der entsprechenden Strategiekonzeption von Winne und Hadwin (1998) erforschten Stahl et al. (2006) die Einflüsse epistemologischer Überzeugungen auf den Lernprozess. Die Stichprobe bestand aus 72 Biologiestudierenden, deren epistemologische Überzeugungen mittels des disziplinübergreifenden *Epistemological Beliefs-Instrument* (EBI, Jacobson & Jehng, 1999), den disziplinspezifischen *Connotative Aspects of Epistemological Beliefs* (CAEB, Stahl & Bromme, 2007) sowie *Domain-specific aspects of Epistemological Beliefs Questionnaire* (DEBQ, Hofer, 2000) erhoben

wurden. Studierende, die der fortgeschritteneren Überzeugung waren, dass Wissen komplex und unsicher ist, setzen bei schweren Aufgaben (Lerninhalt: Genetik) bessere Strukturierungstrategien und nutzen mehrere Wissensquellen als Lernende mit weniger fortgeschrittenen wissenbezogenen Überzeugungen. Darüber hinaus konnten Studierende mit einer relativistischeren Sichtweise auf Wissen die Aufgabenschwierigkeit besser einschätzen, Ziele besser setzen und besser planen als Studierende mit weniger fortgeschrittener Perspektive auf Wissen. Nach Stahl et al. (2006) zeigen diese Ergebnisse, dass Studierende mit fortgeschritteneren epistemologischen Überzeugungen höherwertige Strategien beim Lernen einsetzen.

Unter Einsatz der gleichen Instrumente und Modellannahmen wie Stahl et al. (2006) gelang es Pieschl, Stahl und Bromme (2008), die Ergebnisse der vorherigen Studie anhand einer Stichprobe von 35 Studierenden (21 Geisteswissenschaften, 14 Naturwissenschaften) weitgehend zu replizieren. Auch hier deuteten die Ergebnisse darauf hin, dass sophistiziertere Überzeugungen zu Wissen elaboriertere Lernprozesse verursachen und nicht umgekehrt (ebd.).

Die Zusammenhänge zwischen epistemologischen Überzeugungen und den Softskills zum selbstgesteuerten Lernen untersuchten Barnard et al. (2008). Sie erfassten die epistemologischen Überzeugungen mittels des *Epistemic Belief Inventory* (Hofer, 2004b) und die Softskills zum selbstgesteuerten Lernen mit dem *Online Self-regulated Learning Questionnaire* (OSLQ) (Barnard et al., 2008; Lan, Bremer, Stevens & Mullen, 2004, April) bei 434 amerikanischen Studierenden 18 verschiedener Disziplinen im Rahmen eines akademischen Onlinekurses. Erwartungskonform fanden die Autoren einen signifikanten positiven Zusammenhang zwischen den epistemologischen Überzeugungen und den Softskills zum selbstregulierten Lernen ($r$ = .79, $p$ = .01). Je fortgeschrittener die epistemologischen Überzeugungen der Studierenden, desto besser sind ihre Softskills zum selbstregulierten Lernen. Zwischen dem selbstregulierten Lernen und der akademischen Leistung besteht ebenfalls ein signifikanter, positiver Zusammenhang ($r$ = .13, $p$ = .05).

Paulsen und Feldman (2007) bestätigten die positive Beziehung zwischen fortgeschrittenen epistemologischen Überzeugungen und tiefenorientiertem Lernstrategieeinsatz. Gleichzeitig fanden sie auch empirische Evidenz dafür, dass die Überzeugungen zum Lernen und die epistemologischen Überzeugungen interagieren und gemeinsam den Lernstrategieeinsatz beeinflussen (Paulsen & Feldman, 2007). Paulsen und Feldman (2007) stellten fest, dass die Kombination aus wenig fortgeschrittenen Lernüberzeugungen und wissensbezogenen Überzeugungen die negativen Einflüsse von naiven epistemologischen Überzeugungen auf den Lernprozess verstärkt. Personen, die der Überzeugung waren, dass Wissen sicher und die Lernfähigkeit angeboren ist, nutzen signifikant weniger Elaborationsstrategien und metakognitive Strategien (tiefenorientiert) und mehr Wiederholungsstrategien (oberflächenorientiert).

Bei deutschen Studierenden der Erziehungswissenschaft ($N$ = 210, Alter: $M$ = 22.73, $SD$ = 3.15, ca. 75% weiblich) fanden Schworm und Gruber (2017, S. 152 f.) unter Einsatz des *Fragebogens zur Erfassung epistemischer Überzeugungen* (FEE) (Moschner, Gruber & EPI, 2005) signifikante positive Zusammenhänge zwischen der Überzeugung, dass Lernen gelernt werden

kann und den LIST-Lernstrategien (Wild & Schiefele, 1994) höherer Ordnung (Elaborations-, Organisationsstrategien, kritische Reflexion), den metakognitiven Lernstrategien und der Literatursuche (siehe Tabelle 2.9). Die Überzeugung, dass Wissen unsicher ist und Autoritäten nicht allwissend sind, geht mit signifikant weniger oberflächlichen Lernstrategien (Wiederholungsstrategieeinsatz), kritischer Reflexion und Investition von Aufmerksamkeit und Konzentration einher, aber mit mehr Zeit- und Anstrengungsmanagement, Nutzung der Lernumgebung und Nutzung von Arbeitsgruppen (siehe Tabelle 2.9).

TABELLE 2.9: Bivariate Korrelationen zwischen epistemologischen Überzeugungen und Lernstrategien (vgl. Schworm & Gruber, 2017, S. 153)

| Lernstrategieskalen | Unsicheres Wissen | Lernen zu lernen |
|---|---|---|
| Wiederholung | -.29** | .10 |
| Elaboration | -.12 | .28** |
| Organisation | .08 | .19** |
| Kritische Reflexion | -.19** | .29** |
| Metakognition | .09 | .20** |
| Kooperation | .15* | .13 |
| Anstrengungsmanagement | .24** | .07 |
| Zeitmanagement | .24** | .02 |
| Literatursuche | .06 | .22** |
| Aufmerksamkeit | -.14* | -.04 |
| Lernkontexte | .23** | .15* |

*$p < .05$, **$p < .01$

Auch speziell für Lehramtsstudierende konnte mehrfach gezeigt werden, dass sophistizierte epistemologische Überzeugungen mit tiefenorientierten Lernansätzen einhergehen. Muis (2004) fand starke Hinweise, dass die individuellen epistemologischen Überzeugungen von Lehramtsstudierenden deren Lernstrategien und deren Lernoutcome beeinflussen. In einer Interviewstudie mit 14 Lehramtsstudierenden stellten sich Studierende mit sophistizierten Überzeugungen über unsicheres und integriertes/komplexes Wissen als eher offen für verschiedene Sichtweisen und Zusammenhänge zwischen verschiedenen Perspektiven heraus Bondy et al. (2007). Ravindran, Greene und DeBacker (2005) stellten anhand einer Stichprobe von 101 Lehramtsstudierenden fest, dass sophistizierte epistemologische Überzeugungen verstehensorientierte Lernansätze und eine Lernzielorientierung prädizieren.

Es gibt einige Studien, in denen sich nicht alle sondern vielmehr nur ein oder zwei Überzeugungsdimensionen je nach Aufgabenkontext als einflussreich für Lernsituationen erweisen (z.B. Greene et al., 2010; Hammer & Elby, 2002; Köller, 2001; Muis et al., 2006; Muis & Gierus, 2014; Schommer et al., 1992). Das scheint dafür zu sprechen, dass die epistemologischen Überzeugungen im Lernkontext aktiviert werden. Andere Studien finden entgegen der oben dargestellten Ergebnisse kaum oder keine Belege für die Beziehung zwischen den Dimensionen epistemologischer Überzeugungen und der Neigung zu bestimmten Lernstrategien (Buehl & Alexander, 2004; Neber & Schommer-Aikins, 2002; Schiefele, Streblow,

Ermgassen & Moschner, 2003). Beispielsweise konnten Schiefele et al. (2003) in ihrer Längsschnittstudie mit 285 Studierenden verschiedener Fächer keine Beziehung, weder direkte noch indirekte Zusammenhänge, zwischen den epistemologischen Überzeugungen nach Schommer (1990, 1993a) und dem berichteten Lernstrategieeinsatz sowie der Studienleistung finden.

Trotz einiger solcher widersprüchlicher Ergebnisse und teilweise nur geringen Zusammenhängen zeigt die Befundlage dennoch deutlich, dass die in den Modellen von Schommer-Aikins (2004) und Hofer (2001) angenommenen Beziehungen zwischen epistemologischen Überzeugungen und Lernen sich vielfach bestätigen ließen. Die epistemologischen Überzeugungen von Lernenden – von Studierenden ebenso wie von Schülerinnen und Schülern – nehmen eine bedeutende Rolle in ihrem Lernprozess ein. Es zeigt sich, dass die epistemologischen Überzeugungen als (teilweise unterbewusste) Linse wirken, durch die ein Lernender die Lernaufgabe und das den Lerninhalt betreffende Wissen auffasst (Bromme, 2010, S. 21). Die individuellen Überzeugungen zum Wissen und Wissensprozess scheinen sich auf alle Phasen des selbstregulierten Lernens auszuwirken (Bromme, 2010). Sie beeinflussen die Standards, die ein Lernender für die metakognitive Überwachung und Regulation des Lernprozesses anlegt und regulieren dementsprechend auch die Anpassung des eigenen Lernprozesses an die Aufgabenschwierigkeit, insbesondere über die Lernstrategieauswahl (vgl. Bromme, 2010; Muis, 2007). Über die Standards und Zielsetzungen beim Lernen bestimmen die wissensbezogenen Überzeugungen also die Auseinandersetzung mit der Lernaufgabe, die Tiefe der Verarbeitung und die Überwachung des eigenen Verstehens (Hofer & Pintrich, 1997).

Allgemein zeigt der Forschungsüberblick, dass naivere, weniger sophistizierte Überzeugungen, zum Beispiel die Überzeugung, dass Wissen einfach und unveränderlich ist, mit weniger adaptiven akademischen Kognitionen und schlechterer Leistung verbunden sind (Bråten & Olaussen, 2005, S. 144). Ein sophistizierteres Überzeugungssystem kann als generell förderlicherer für die verschiedensten Lernhandlungen, den Lernprozess selbst und das Lernoutcome betrachtet werden, im Speziellen bezüglich der Auswahl und des Einsatzes von tiefenorientierten Lernstrategien. Dies gilt für Lernende verschiedenster Altersgruppen und in den verschiedensten Lernkontexten und natürlich auch für Lehramtsstudierende (Schraw & Sinatra, 2004).

### 2.8.3 Epistemologische Überzeugungen beim Lehren

Die individuellen wissensbezogenen Überzeugungen beeinflussen nicht nur den Lernprozess sondern auch den Lehrprozess (siehe Abb. 2.8, S. 87, Hofer, 2001). Die epistemologischen Überzeugungen eines Lehrenden, wie beispielsweise von Hochschuldozentinnen und -dozenten (z.B. Strømsø & Bråten, 2011b) oder Lehrerinnen und Lehrern, beeinflussen, „wie der Lehrende unterrichtliche Situationen auffasst, welche Handlungsmöglichkeiten er in Betracht zieht und welche er letztlich auswählt, wie er Effekte seines Eingreifens bewertet und wie er nachträglich sein gesamtes Handeln begründet oder rechtfertigt“ (Wahl, 1979, S. 209).

Einige Studien konnten den Einfluss von epistemologischen Überzeugungen sowie den Vorstellungen zu Lernen und Lehren von Lehrerinnen und Lehrern auf ihr unterrichtliches Handeln (Lehrtätigkeit, Lehrstrategien) und ihre Erwartungen an die Schülerinnen und Schüler bestätigen (z.B. Brickhouse, 1989; Brownlee et al., 2011; Buehl & Fives, 2009; Hammrich, 1998; Hashweh, 1996; Muis & Franco, 2010; Tsai, 2002, 2006) und wiederum eine Verbindung zwischen unterrichtlichem Handeln und den epistemologischen Überzeugungen der Schülerinnen und Schüler identifizieren (z.B. Hofer, 1999; Muis & Franco, 2010; Schoenfeld, 1988; Stodolsky, 1985).

Die epistemologischen Überzeugungen von Lehrerinnen und Lehrern wirken sich auf ihre Lehrhandlungen, die Unterrichtsstrategien und ihre Erwartungen an die Schülerinnen und Schüler aus (vgl. Brownlee et al., 2011; Buehl & Fives, 2009). Daniels und Shumow (2003) konnte zeigen, dass die epistemologischen Überzeugungen von Lehrkräften ihre Handlungen und wiederum die Lernumgebung der Schülerinnen und Schüler beeinflussen und die Lernumgebung sich auf die Handlungen der Schülerinnen und Schüler auswirkt. Auf Basis ihres Reviews schlussfolgerten die Autoren, dass Schülerinnen und Schüler in konstruktivistisch orientierten Lernumgebungen (schülerorientiert) über höhere Motivation, geringeren Stress sowie höhere Problemlösefähigkeiten und Sprachfähigkeiten verfügen im Vergleich zu Kindern in klassisch lehrerorientierten Lernumgebungen (Muis & Franco, 2010, S. 437). Seidel, Schwindt, Rimmele und Prenzel (2008) zeigten Zusammenhänge zwischen weniger fortgeschrittenen epistemologischen Überzeugungen bei Lehrkräften und traditionellen lehr- und lerntheoretischen Überzeugungen (Lehrerorientierung). Wobei Lehrkräfte mit konstruktivistischen Ansichten, d.h. solche die annehmen, dass die Schülerinnen und Schüler in der Lage sind, selbst Wissen zu konstruieren, über ein breiteres Repertoire an Lehrstrategien verfügen und auch eher Lehrstrategien zur Förderung von *Conceptual Change* einsetzen, als Lehrkräfte mit traditionellen Sichtweisen im Sinne von Schülerinnen und Schülern als passive Rezipienten von Wissen (Hashweh, 1996). Tsai (2006) fand mittels einer Interviewstudie Kohärenzen zwischen den epistemologischen Überzeugungen von Lehrkräften zu naturwissenschaftlichem Wissen und ihren Lehrhandlungen. Lehrkräfte, die der Ansicht waren, dass das naturwissenschaftliche Wissen sicher ist und einfach nur entdeckt werden muss, konzentrierten sich auf die Testergebnisse der Schülerinnen und Schüler, verbrachten mehr Zeit mit Lehrervorträgen sowie geleiteten Übungsaufgaben und in den Unterricht integrierte Tests. Lehrkräfte dagegen, die der Ansicht waren, dass das naturwissenschaftliche Wissen von Wissenschaftlern konstruiert wird, veränderlich ist und

dem gesellschaftlichen Dialog bei der Begründung von Wissensansprüchen Bedeutung beimaßen, konzentrierten sich in ihrem Unterrichten mehr auf die Anwendung des Gelernten der Schülerinnen und Schüler sowie deren Verstehen der Inhalte. Auch verbrachten sie mehr Unterrichtszeit mit der schülerorientierten Entdeckung der Inhalte und interaktiven Diskussionen. Auf Basis der Ergebnisse einer längsschnittlichen Studie zu den epistemologischen Überzeugungen von Schülerinnen und Schülern über Mathematik, die die Auswertung von Unterrichtsprozessen, Beobachtungen und Interviews umfassten, wurde wiederum der bedeutende Einfluss von im Unterricht angewandten Lehrstrategien auf die epistemologischen Überzeugungen der Schülerinnen und Schüler gezeigt (Schoenfeld, 1988). Hofer (1999) verglich die akademische Leistung, Motivation und Lernstrategien von Studierenden in zwei unterschiedlichen Lernumgebungen (im Fach Mathematik). In der einen Lernumgebung wurden klassische Lehrmethoden angewandt. Der Lehrende nutzte hier einen Standardtext, der die Inhalte sequenziell enthielt sowie Vorträge und Demonstrationen von Problemstellungen. In der zweiten Lernumgebung wurde ein sozial-konstruktivistischer Lehransatz umgesetzt. Der Schwerpunkt lag auf kollaborativem Lernen und der aktiven Arbeit der Lernenden mittels komplexer, in Kontext eingebetteter Problemstellungen mit mehreren möglichen Lösungen. Die Lernenden der konstruktivistischen Lernumgebung verfügten am Ende des Kurses über sophistiziertere Überzeugungen zu Wissen als die Lernenden des anderen Kurses. Sophistiziertere epistemologische Überzeugungen auf Lernerseite gingen außerdem mit höherer intrinsischer Motivation, Selbstwirksamkeit, Selbstregulation und besseren Noten einher (ebd.).

Kang (2008) fand anhand einer Stichprobe von 23 Lehramtsstudierenden der Naturwissenschaft heraus, dass die epistemologischen Überzeugungen von angehenden Lehrerinnen und Lehrern sowohl mit den von ihnen gesteckten Lehrzielen zusammenhängen, als auch mit den zur Zielerreichung konkret im Unterricht umgesetzten pädagogischen Handlungen. Lehramtsstudierende mit konstruktivistischer ausgeprägten epistemologischen Überzeugungen setzten sich konstruktivistischere Lehrziele. Yadav und Koehler (2007) konnten übereinstimmend damit zeigen, dass die epistemologischen Überzeugungen von angehenden Lehrerinnen und Lehrern in einer Verbindung damit stehen, was diese als effektive Didaktik bezüglich des Lese- und Schreibunterrichts betrachten. Angehende Lehrkräfte mit fortgeschritteneren wissensbezogenen Überzeugungen ermutigen die Kinder eher zum selbstständigen Finden von Antworten und sehen Fehler eher als Lerngelegenheiten, während angehende Lehrkräfte, die der Überzeugung sind, dass Wissen einfach ist, sich in ihren pädagogischen Handlungen eher auf das Verhaltensmanagement (die Klassenführung) fokussieren.

Muis und Franco (2010) konnten anhand einer Stichprobe von 55 amerikanischen Grundschullehrkäften (48 weiblich, 7 männlich, Alter: $M$ = 38.84, $SD$ = 10.50) und 131 Schülerinnen und Schülern der Jahrgangsstufe vier und fünf von 26 verschiedenen Schulen die in

Schommer-Aikins (2004) *Embedded Systemic Model* angenommenen Pfade zwischen Lehrerüberzeugungen (zu Wissen und sowie Lehren und Lernen) und der Leistung und den Überzeugungen der Schülerinnen und Schüler empirisch bestätigen. Sophistizierte epistemologische Überzeugungen auf Lehrerseite gingen sophistizierten Überzeugungen auf Schülerseite und mit höherer Matheleistung einher und wiesen darüber hinaus eine negative Verbindung mit wenig fortgeschrittenen Überzeugungen in einfaches und sicheres Mathematikwissen auf Schülerseite auf. Weniger sophistizierte Überzeugungen über die Sicherheit und Einfachheit des Wissens auf Schülerseite gingen mit Leistungszielorientierung einher, wiesen aber eine negative Verbindung zu Lernzielorientierung auf. Die Überzeugung auf Schülerseite, dass Lernen Anstrengung erfordert, sagte Lernzielorientierung positiv und Leistungszielorientierung negativ vorher. Selbstwirksamkeit stellte sich als Mediator zwischen der Zielorientierung und der Leistung heraus. Leistungszielorientierung prädizierte eine niedrigere Selbstwirksamkeit der Schülerinnen und Schüler, Lernzielorientierung dagegen eine höhere, während die Selbstwirksamkeit wiederum die Leistung prädizierte.

Diese Ergebnisse zum Einfluss der epistemologischen Überzeugungen von Lehrenden stützen einerseits die in Hofers (2001) Modell postulierten Verbindung zwischen den epistemologischen Überzeugungen der Lehrenden und dem pädagogischen Handeln im Unterricht (siehe Abb. 2.8, S. 87). Darüber hinaus bekräftigen sie auch die im Rahmen des TIDE Modells (siehe Abb. 2.6, S. 76) postulierten Annahmen darüber, dass die Interaktion mit der Umwelt eine bedeutende Rolle bei der Entwicklung, Veränderung sowie dem Ausdruck von individuellen Überzeugungen zu Wissen und dem Wissensprozess spielt (Muis & Franco, 2010, S. 452). Individuen konstruieren und deuten ihre Erfahrungen aktiv, während eine Weiterentwicklung und Veränderung der wissensbezogenen Überzeugungen als Funktion der Interaktion des Individuums mit der sozialen Umwelt stattfindet (vgl. Bendixen & Rule, 2004; Hofer & Pintrich, 1997; Muis & Franco, 2010).

Aufgrund der bedeutenden Rolle, die den individuellen epistemologischen Überzeugungen beim Lernen und Lehren zukommt, ist es wenig verwunderlich, dass zahlreiche Verfahren zur Erfassung derselbigen existieren. Denn um die wissensbezogenen Überzeugungen erforschen zu können und letztendlich zu verstehen, müssen sie zunächst empirisch zugänglich gemacht werden. Den verschiedenen bislang bestehenden Erhebungsverfahren epistemologischer Überzeugungen ist das folgende Kapitel gewidmet.

## 2.9 Erhebung epistemologischer Überzeugungen

Wie aus der Beschreibung der unterschiedlichen Konzeptionen epistemologischer Überzeugungen bereits hervorgeht (siehe Kapitel 2.2, S. 25), gehen mit den Ansätzen epistemologischer Überzeugungen unterschiedliche Definitionen des Konstrukts einher. Unter anderem aus diesen divergierenden theoretischen Annahmen ergibt sich, dass verschiedene Methoden der Erhebung eingesetzt wurden und werden (Mason, 2016). Gleichzeitig erfordern auch die „unterschiedlichen Zielsetzungen der beispielweise hypothesenprüfenden, hypothesengenerierenden oder rekonstruierenden Forschungsbemühungen“ (Zinn, 2013, S. 108) den Einsatz unterschiedlicher qualitativer oder quantitativer Erhebungsmethoden.

In diesem Kapitel wird zunächst ein kurzer Überblick darüber gegeben, welche qualitativen Methoden der Erfassung epistemologischer Überzeugungen in der Vergangenheit eingesetzt wurden und bis heute eigesetzt werden. Da im Zentrum dieser Arbeit jedoch die Entwicklung eines quantitativen Erhebungsinstruments steht, wird im Anschluss ausführlicher auf die quantitativen disziplinübergreifenden und -spezifischen Erhebungsinstrumente eingegangen. Darauf folgend werden die Schwierigkeiten und Stärken der bestehenden quantitativen Erhebungsinstrumente herausgearbeitet.

### 2.9.1 Qualitative Erhebung

Das in der ersten Forschungsphase zu epistemologischen Überzeugungen vorherrschende Forschungsinteresse führte dazu, dass hauptsächlich Interviewverfahren zur Erhebung eingesetzt wurden. Entwicklungspsychologen betrachteten epistemologisches Denken als integrierten, disziplinübergreifenden Entwicklungsverlauf der kognitiven Strukturen – die Repräsentationen zu Wissen enthalten – über verschiedene, qualitativ unterschiedliche Stufen hinweg (Mason, 2016). Das individuelle Denkvermögen und die Repräsentationen zu Wissen werden als eng miteinander verbunden angesehen (Mason, 2016). Dementsprechend wurden Intensivinterviews (z.B. Baxter Magolda, 1992; Belenky et al., 1986; Perry, 1970) eingesetzt oder aber Dilemmainterviews, d.h. Interviews auf der Basis von komplexen Problemstellungen, die durch die Interviewten aufzulösen und zu begründen waren (z.B. King & Kitchener, 1994; Kuhn, 1991). Die umfassenden Antworten der Interviewten enthielten ihre Reflexionsprozesse und Schlussfolgerungen sowie deren zugrundeliegenden Annahmen zu Wissen, dessen Grenzen, Sicherheit und die angelegten Kriterien für Wissensurteile. Diese Antworten wurden transkribiert, inhaltsanalytisch ausgewertet und die Interviewten so einer ihrem epistemologischen Denken entsprechenden Stufe der epistemologischen Entwicklung (z.B. *pre-reflective thinking*, *quasi-reflective thinking*, *reflective thinking*, King & Kitchener, 1994) zugeordnet. Es hat sich allerdings in Bezug auf die Dilemmainterviews gezeigt, dass ein Individuum über verschiedene Problemstellungen hinweg unterschiedlichen Stufen der epistemologischen Entwicklung zuzuordnen wäre (Mason, 2016, S. 377). Hierin liegt auch einer der Kritikpunkte dieses Verfahrens, denn dem theoretischen Anspruch der Forschenden entsprechend müsste es möglich sein, ein Individuum genau einer Stufe eindeutig zuzuordnen; liegt doch die Annahme von epistemologischem Denken als Entwicklungsverlauf von speziellen kognitiven Strukturen, die disziplinübergreifendene Repräsentationen

zu Wissen enthalten, zugrunde (ebd.). King und Kitchener (2002, S. 57) selbst vermuten außerdem, dass ihre Interviews die kognitiven Fähigkeiten der Befragten unterschätzen könnten, da diese unvorbereitet und ohne Zeit zum Nachdenken antworten mussten. Eine Stärke dieser Verfahren ist hingegen, dass die Befragten in den Interviews die Möglichkeit bekommen, ihre Gedankengänge und Schlussfolgerungen sehr ausführlich darzulegen, was dem Forschenden ausführliche Einblicke in die individuellen Annahmen zu Wissen und dem Wissensprozess zur vorgegebenen Problemstellung gewährt (Mason, 2016). Problematisch ist jedoch, dass die Erhebung mittels Interviews sehr zeit- und kostenintensiv ist, da die Interviews nur von speziell dafür geschulten Personen und immer nur einzeln durchgeführt werden können sowie in der Durchführung und Auswertung sehr aufwändig sind (ebd.).

Neben dem Intensiv- und Dilemmainterview gibt es zahlreiche weitere qualitative Erhebungsverfahren, die zur Erfassung epistemologischer Überzeugungen eingesetzt werden. Das umfasst das Dialog-Konsens-Verfahren (vgl. z.B. Scheele & Schreier, 1994), die Strukturlegeverfahren (vgl. z.B. Christmann & Groeben, 1999; Schlee, 1988), das Netzinterview (vgl. z.B. Müller, 2009), die Methode lauten Denkens (vgl. z.B. Kardash & Noel, 2000; Schommer et al., 1992; Trautwein et al., 2004) und die nicht teilnehmende Beobachtung (vgl. z.B. Lederman & Zeidler, 1987; Seidel & Prenzel, 2006; Seifried, 2006).

Da im Zentrum dieser Arbeit die Entwicklung eines quantitativen Erhebungsinstruments und ein hypothesenprüfendes Forschungsinteresse stehen, wird an dieser Stelle auf eine Darstellung der verschiedenen qualitativen Verfahren verzichtet. Ein tabellarischer Überblick zu den quantitativen und qualitativen Instrumenten zur Erhebung allgemeiner epistemologischer Überzeugungen ist bei Duell und Schommer-Aikins (2001) zu finden.

### 2.9.2 Quantitative Erhebung

Vorreiter der quantitativen Erhebung von epistemologischen Überzeugungen ist M. P. Ryan (1984), der die epistemologischen Überzeugungen erfasste, indem er die Versuchspersonen Einstellungsstatements bewerten ließ. Mit der Wende hin zur Annahme expliziter Mehrdimensionalität der epistemologischen Überzeugungen (Schommer, 1990) ging ein genereller Umschwung in Richtung quantitativer Erfassung epistemologischer Überzeugungen mittels Fragebögen einher, die bis heute vorherrschend ist. Auf Basis von theoretisch angenommenen Dimensionen wird ein Itempool entwickelt und in einer Stichprobe eingesetzt. Die erfassten Daten werden faktorenanalytisch bezüglich ihrer Dimensionalität geprüft und die Itemauswahl optimiert. Das resultierende Instrument spiegelt dann im besten Fall das zugrundeliegende theoretische Modell wider oder aber einzelne Teilaspekte davon. Die Dimensionalität des Instruments kann im Folgenden anhand der Daten neuer Stichproben überprüft werden. Diese Fragebögen, die in der zweiten Forschungsphase entstanden, waren allerdings zunächst alle auf die Erfassung der allgemeinen, disziplinübergreifenden Überzeugungen zu Wissen ausgelegt. Der erste und zugleich bekannteste Fragebogen, der die mehrdimensionale Erhebung epistemologischer Überzeugungen von Studierenden ermöglicht, ist der aus dem angloamerikanischen Raum stammende *Schommer Epistemological Questionnaire* (SEQ, Schommer, 1990). Der SEQ und die zugrundeliegende theoretische

Konzeption (vgl. Kapitel 2.2.4) von Schommer (1990) stellen die Basis zahlreicher Fragebogenentwicklungen dar (siehe Abb. 2.13, S. 116). Beispielsweise setzten Kardash und Scholes (1996), Kardash und Noel (2000) und Mori (1999) bei US-amerikanischen Studierendenstichproben gekürzte Versionen des SEQ mit siebenstufigen bzw. sechsstufigen Ratingskalen statt einer fünfstufigen Ratingskala wie im Original ein (vgl. Buehl, 2008, S. 71 f.). Es existieren bereits Versionen des SEQ für Schülerinnen und Schüler in den verschiedensten Sprachen, wie koreanisch (Youn, Yang & Choi, 2001), chinesisch (Lin, 2002), italienisch (Mason, 2000), spanisch (Cano, 2005) und englisch (Qian & Alvermann, 1995; Schommer-Aikins et al., 2005, 2000) sowie Versionen für Studierende in deutsch (Müller, Heidemann & Rebmann, 2006), norwegisch (Bråten & Olaussen, 2005), chinesisch (Chan & Elliott, 2000) und holländisch (Clarebout et al., 2001). Einige Forschende nutzen Schommers SEQ in Kombination mit anderen Instrumenten und Konzeptionen als Basis für ganz eigene Instrumententwicklungen. So entstand beispielsweise das ebenfalls sehr bekannte, englischsprachige *Epistemic Belief Inventory* (EBI, Schraw et al., 1995) sowie der ebenfalls englische *Jehng's Epistemological Questionnaire* (JEQ Jehng et al., 1993) (siehe Abb. 2.13, S. 116). Auf Basis des JEQ wurden zum Beispiel der *Beliefs about Learning Questionnaire* (BLQ) für englischsprachige Studierende (Cole & Goetz, 2000) oder die in sowohl englischer als auch koreanischer Form vorliegende *Epistemological belief scale* (Youn, 2000) für Studierende entwickelt. Der EBI wurde weiterentwickelt (siehe z.B. Bendixen et al., 1998; Jacobson & Jehng, 1999; Nussbaum & Bendixen, 2003; Schraw et al., 2002) und wird bis heute eingesetzt (z.B. Blignaut et al., 2010; Mokwinski, 2011; Pieschl, Stahl & Bromme, 2006 und EBI-AM Peter, 2015). Wood und Kardash (2002) nutzen für die Erhebung der epistemologischen Überzeugungen von Studierenden eine Kombination der Items des SEQ und des JEQ. Von Galloti et al. (1999) stammt der ebenfalls englischsprachige und disziplinübergreifende *Attitudes Toward Thinking and Learning Survey* (ATTLS).

Ganz ungeachtet der mehrdimensionalen Erhebungsinstrumente entstanden auf Basis der Kritik an der Erhebung mittels Interviews allerdings auch entwicklungspsychologisch ausgerichtete, ökonomischere Erhebungsverfahren zur Ermittlung der Stufe epistemologischer Entwicklung. Im Rahmen des Erhebungsverfahrens von Wood und Kardash (2002), werden die Probanden dazu aufgefordert, verschiedene Statements, die unterschiedliche Stufen der epistemologischen Entwicklung widerspiegeln, zu lesen. Danach schreiben die Probanden selbst kurze Statements und schätzen im Anschluss auf einer Likertskala die Passung der vorgegebenen Statements zu den eigenen ein (Mason, 2016, S. 377). Mansfield und Clinchy (2002) nutzten ein Verfahren mit Vignetten, Text, Bildern und ergänzenden Interviews zur Erfassung der epistemologischen Entwicklung von Schülerinnen und Schülern. Auch Kuhn und Pearsall (2000) entwickelten auf Basis ihrer entwicklungspsychologischen Konzeption ein *paper-and-pencil* Erhebungsverfahren zur Erhebung des Levels epistemologischer Entwicklung. Allerdings bezieht sich dieses Instrument auf Urteilsbereiche, wie Ästhetik und persönliche Werte (Mason, 2016, S. 377).

Während sich unter Bildungsforschern und in der pädagogischen Psychologie nach und nach die Ansicht durchsetzte, dass die epistemologischen Überzeugungen von Individuen

| Bezeichnung | Autoren | theoretische Basis | Faktoren (Itemanzahl, α) | Nr. Item | Itemart | Zielgruppe |
|---|---|---|---|---|---|---|
| Schommer Epistemological Questionnaire (SEQ), auch genannt: Epistemological Belief Questionnaire (EBQ) | Schommer (1990) | Schommer (1990) | simple knowledge (seek single answer, 11; avoid integration, 8), certain knowledge (avoid ambiguity, 5; knowledge is certain, 6), innate ability (can't learn how to learn, 5; success is unrelated to hard work, 4; ability to learn is innate, 4), quick learning (learning is quick, 5; learn first time, 3; concentrated effort is a waste of time 2), [*omniscient authority (don't criticize authority, 6; depend on authority, 6)]* (α: n.a.) (Varianzaufklärung: 55,2%) | 63, 12 subsets | 5-stufige Ratingskala | Studierende |
| Jehng's Epistemological Questionnaire (JEQ) | Jehng et al. (1993) | Schommer (1990) und Spiros (1989) Fragebögen | certainty of knowledge, omniscient authority, orderly process, innate ability, quick learning (α: .42-.59) | 51 | 7-stufige Ratingskala | Studierende |
| Epistemic Belief Inventory (EBI) | Schraw et al. (1995); Schraw, Bendixen & Dunkle (2002) | Schommer (1990) | fixed ability (Study I: 5, .87; Study II: 4, .84), certain knowledge (Study I: 3, .76; Study II: 4, .76), omniscient authority (Study I: 3, .76; Study II: 3, .71), simple knowlegde (Study I: 2, .67; Study II: 2, .63), quick learning (Study I: 3, .74; Study II: 3, .73) (Varianzaufklärung: 60%-64%) | 32 | 5-stufige Ratingskala | Studierende |

ABBILDUNG 2.13: Bedeutende englischsprachige Fragebögen zur mehrdimensionalen, disziplinübergreifenden Erfassung epistemologischer Überzeugungen (überarbeitet und in eigener Darstellung nach Buehl, 2008, S. 70 ff.)

nicht nur mehrdimensional sondern auch abhängig von der betreffende Disziplin sind (vgl. Muis et al., 2006), mehrten sich die Fragebogenkonstruktionen, die diesen Aspekt berücksichtigten und dementsprechend sowohl mehrdimensional als auch auf eine oder mehrere spezifische akademische Disziplinen ausgerichtet sind. Einige bedeutende englischsprachige, disziplinspezifische Fragebögen sind in Abbildung 2.14 auf Seite 118 überblicksartig dargestellt.

Unter Rückgriff auf Items von Lampert (1990) und Schoenfeld (1992), entwickelte Hofer (1999) spezifisch für die Erfassung der Überzeugungen über die Einfachheit mathematischen Wissens bei Studierendenden eine Kurzskala mit sechs Items. Buehl et al. (2002) entwickelten auf Basis des SEQ den englischsprachigen *Domain-Specific Belief Questionnaire* (DSBQ) für die Erfassung der epistemologischen Überzeugungen von Studierenden bezüglich der Disziplinen Geschichte und Mathematik. Von Hofer (2000) wurde der ebenfalls englische und auf Studierende ausgerichtete *Discipline-Focused Epistemological Beliefs Questionnaire* (DEBQ) mit 27 Items konstruiert. Die theoretische Grundlage des DEBQ besteht in der Konzeptionierung von Hofer und Pintrich (1997) (siehe Kap. 2.2.5, ab S. 43). Damit unterscheidet sich der DEBQ ganz fundamental von allen zuvor genannten, mehrdimensionalen Erhebungsinstrumenten. Der DEBQ deckt nur die Kerndimensionen epistemologischer Überzeugungen ab und beinhaltet, anders als der SEQ und verwandte Instrumente, keine Skalen zu Überzeugungen zum Lernen. Auf der Grundlage der Konzeption von Hofer und Pintrich (1997), entwickelten Bråten und Olaussen (2005) den *Internet-specific Epistemological Questionnaire* (ISEQ) in norwegischer Sprache für Studierende. Eine parallele englischsprachige Version des ISEQ stammt von Bråten und Strømsø (2004). Auch der *Epistemology Survey* für Studierende in englischer und arabischer Version von Karabenick (2005) basiert auf der Konzeption von Hofer und Pintrich (1997). Unter Rückgriff sowohl auf Items des DSBQ (Buehl et al., 2002) zur Isoliertheit von Wissen als auch auf Items des DEBQ (Hofer, 2000) zur Einfacheit des Wissen sowie Autorität als Wissensquelle, erhoben Buehl und Alexander (2004, 2005) die Überzeugungen zum geschichtlichen und mathematischen Wissen bei Studierenden. Es liegen einige englischsprachige Fragebögen zur Erfassung der individuellen Überzeugungen zum Wissen und Wissensprozess in Physik vor, wie der *Epistemological Beliefs Assessment for Physical Science* (EBAPS, Elby, 2001), der *Views about Science Survey* (VASS, Halloun & Hestenes, 1998) und der *Maryland Physics Expectations* (MPEX, Redish, 1997). Bei dem *Colorado Learning Attitudes about Science Survey* (CLASS, Barbera, Perkins, Adams & Wieman, 2008) und *Views of Nature of Science* (VNOS, Lederman, Abd-El-Khalick, Bell & Schwartz, 2002) handelt es sich ebenfalls um englischsprachige Fragebögen, die sich auf Überzeugungen zum naturwissenschaftlichen Wissen und Lernen beziehen.

| Bezeichnung | Autoren | theoretische Basis | Faktoren (Itemanzahl, α) | Nr. Item | Itemart | Disziplin | Zielgruppe |
|---|---|---|---|---|---|---|---|
| Discipline-Focused Epistemological Belief Questionnaire DEBQ (teilw. DFEBQ) | Hofer (2000) | Hofer & Pintrich (1997) | certain/simple knowledge (8, psychology: .74; science: .81), source of knowledge: authority (4, psychology: .51; science: .64, justification for knowing: personal (4, psychology: .56; science: .61, attainability of truth (2, psychology: .60; science: .75) | 27 | 7-stufige Ratingskala | Psychologie, Naturwissenschaft | Studierende |
| | Elder (2002) | Hofer & Pintrich (1997) | changing nature of science (9, .67), role of experiments and source of knowledge (6, .52), source of knowledge (4, .64) | 25 | 5-stufige Ratingskala | Naturwissenschaft | Schülerinnen und Schüler |
| Certain Epistemological Belief Scale | Conley et al. (2004) | adaptiert von Elder (2002), Hofer & Pintrich (1997) | certainty of knowledge (6, T1: .78, T2: .79), source of knowledge (5, T1: .81, T2: .82), justification of knowledge (9, T1: .65, T2: .76), development of knowledge (6, T1: .57, T2: .66) | 26 | 5-stufige Ratingskala | Naturwissenschaft | Schülerinnen und Schüler |
| Domain-Specific Belief Questionnaire (DSBQ) | Buehl et al. (2002) | Schommer (1990) | integration of information and problem solving in mathematics (6, Study II: 70, Study III: 69), integration of information and problem solving in history (6, Study II: 75, Study III: 65), need for effort in mathematics (5, Study II: 68, Study III: 72), need for effort in history (5, Study II: 61, Study III: 58) (Varianzaufklärung: 33,3%) | 22 | 10-stufige Ratingskala | Mathematik, Geschichte | Studierende |
| | Buehl & Alexander (2005) | Items von DSBQ (zu isolation of knowlegde) und DEBQ (zu certainty und authority) | isolation of history knowledge beliefs (.77), certainty of history knowledge beliefs (.70), authority as the source of history knowledge (.74), isolation of mathematics knowledge beliefs (.77), certainty of mathematics knowledge beliefs (.70), authority as the source of mathematics knowledge (.74) | n.a. | 10-stufige Ratingskala | Mathematik, Geschichte | Studierende |

ABBILDUNG 2.14: Bedeutende englischsprachige Fragebögen zur mehrdimensionalen, disziplinspezifischen Erfassung epistemologischer Überzeugungen (überarbeitet und in eigener Darstellung nach Buehl, 2008, S. 70 ff.)

Da die ersten Fragebögen zur Erfassung epistemologischer Überzeugungen alle aus der US-amerikanischen Forschung kommen, handelt es sich bei den existierenden deutschen Fragebögen häufig um direkt übersetzte oder übersetzte und modifizierte Versionen der bestehenden englischsprachigen Fragebögen. Aus Abbildung 2.15 auf Seite 120 ist zu entnehmen, dass es einige deutschsprachige mehrdimensionale, disziplinübergreifende Versionen des SEQ und auch des EBI gibt, beispielsweise von Gerber (2004), Schiefele und Moschner (1997), Müller et al. (2006) und Pieschl et al. (2006) sowie Paechter, Müller, Manhal und Rebmann (2007). Im deutschsprachigen Raum haben Gruber und Moschner (2003) unter Rückgriff auf Aspekte von Belenky et al. (1986), Baxter Magolda (1992), Schommer (1990) und King und Kitchener (1994) einen eigenständigen Ansatz epistemologischer Überzeugungen entwickelt, der die drei Komponenten Begabung, Lernen und Wissen beinhaltet. Auf Basis dieser Konzeption entwickelten sie einen Fragebogen, der in der Validierung noch sieben Dimensionen umfasste (siehe FEE, Abb. 2.15, S. 120).

Die wichtigsten deutschen disziplinspezifischen Fragebögen sind in Abbildung 2.16 (S. 121) aufgeführt. Der *Connotative Aspects of Epistemological Beliefs* CAEB von Bartholome, Stahl, Pieschl und Bromme (2006) und Stahl und Bromme (2007) verfügt über die beiden Dimensionen *Textur* und *Variabilität*, die sich faktorenanalytisch aus den ursprünglich angenommenen 24 Adjektivpaaren zu den Dimensionen *Struktur*, *Sicherheit* und *Quelle des Wissens* ergeben haben. Der *Fragebogen zur Erfassung des Entwicklungsniveaus epistemologischer Überzeugungen* (FREE, Krettenauer, 2005) dient zur Erfassung des Entwicklungsniveaus epistemologischer Überzeugungen. Die Befragten werden im Rahmen eines halbstandardisierten Interviews mit kontroversen Aussagen zu den Disziplinen Geisteswissenschaften, Naturwissenschaften und Sozialwissenschaften konfrontiert, zu denen sie Stellung nehmen müssen. Die Befragten bewerten die vorgelegten Konflikte zunächst auf einer sechsstufigen Likert-Skala und geben zusätzlich noch an, welche der Meinungen der eigenen am besten entspricht. Über ein Verrechnungsverfahren werden die Befragten im Anschluss einem der drei Entwicklungsniveaus Absolutismus, Relativismus oder Postrelativismus zugeordnet.

Eine theoriespezifische Erfassung epistemologischer Überzeugungen im naturwissenschaftlichen Wissensbereich ermöglicht der Fragebogen von Trautwein et al. (2004). Trautwein et al. (2004) setzten darüber hinaus auch eine Kurzskala mit nur drei Items zur Erfassung der individuellen studienfachspezifischen epistemologischen Überzeugungen ein. Die drei Items beziehen sich ausschließlich auf die Überzeugung zur Veränderlichkeit des Wissens im Studienfach. Urhahne und Hopf (2004) nutzen eine übersetzte Version des Fragebogens von Conley et al. (2004), um die Überzeugungen zur Sicherheit, Quelle, Rechtfertigung und Entwicklung zum naturwissenschaftlichen Wissen bei Schülerinnen und Schülern zu erheben. Diesen Fragebogen setzte Urhahne und Harms (2006) auch bei Studierenden der Biologie ein. Eine übersetzte Version des DEBQ wurde von Pieschl et al. (2006) bei Studierenden eingesetzt, bei der faktorenanalytisch die Fakoren *Sicherheit* und *Zugänglichkeit* identifiziert wurden (siehe Abb. 2.16, S. 121).

| Bezeichnung | Autoren | theoretische Basis | Faktoren (Itemanzahl, α) | Nr. Items | Itemart | Zielgruppe |
|---|---|---|---|---|---|---|
| Fragebogen zur Erfassung epistemischer Überzeugungen (FEE) | Moschner & Gruber (2005) | Gruber & Moschner (2003); SEQ, Schommer (1990) | Reflexivität des Wissens (n.a.), Absolutes Wissen (12, .82), Wert von Wissen (5, .79), Lernen lernen (6, .76), Relativität von Wissen, geschlechtsspezifische Wissenszugänge (10, .86), kulturspezifische Wissenszugänge (7, .87), soziale Komponente des Wissens (5, .74) | n.a. | 6-stufige Ratingskala | Studierende |
| stark modifizierte und ergänzte Übersetzung des SEQ | Schiefele & Moschner (1997); Schiefele, Moschner & Husstegge (2002); Schiefele, Streblow & Brinkmann (2007) | SEQ, Schommer (1990) | Objektivität des Wissens (12, .90), Unbeeinflussbarkeit der Lernfähigkeit (5, .69), Komplexität des Lernprozesses (7, .72), Eindeutigkeit der Wissensvermittlung (9, .79), Faktenorientierung bei Prüfung (7, .76) | 40 | 6-stufige Ratingskala | Studierende |
| bearbeitete Übersetzung des SEQ | Gerber (2004) | SEQ, Schommer (1990) | Sicherheit des Wissens (8, .72), Komplexität des Wissen (10, .79), Wissensquelle (7, .76), Wissenbegründung (10, .87) (Varianzaufklärung: 39%) (verkürzte Version für Kinder, Varianzaufklärung: 40%) | 35 (27) | 4-stufige Ratingskala | Eltern (Kinder) |
| SEQ_d | Paechter, Müller, Manhal & Rebmann (2007) | SEQ, Schommer (1990) | Struktur, Sicherheit (α: .63 bis .70) (Varianzaufklärung: 23,97%) | 52 | Ratingskala | Studierende |
| EBI_d | Müller, Heidemann & Rebmann (2006) | übersetzte Version des EBI von Schraw et al. (1995); Schommer (1990) | Struktur, Sicherheit, Kontrolle, Geschwindigkeit, Quelle (α: .55 bis .69) (Varianzaufklärung: 54,06%) | 28 | 5-stufige Ratingskala | Studierende |
| deutsche Übersetzungen des EBI | Pieschl, Stahl & Bromme (2006) | EBI von Jacobson & Jehng (1999); Schommer (1990) | resultierende EBI Dimensionen: Sicherheit (7, .76), Einfachheit (7, .69) (Varianzaufklärung: 42 %) | 14 | Ratingskala | Studierende |
| Relativismus-Skala | Trautwein, Lüdke & Beyer (2004) | Items von Hofer (2000); Schommer (1990); Schiefele et al. (2002) | Struktur des Wissens (3, .67) | 3 | 4-stufige Ratingskala | Schülerinnen und Schüler |

ABBILDUNG 2.15: Bedeutende deutschsprachige Fragebögen zur mehrdimensionalen, disziplinübergreifenden Erfassung epistemologischer Überzeugungen

| Bezeichnung | Autoren | theoretische Basis | Faktoren (Itemanzahl, α) | Nr. Items | Itemart | Disziplin | Zielgruppe |
|---|---|---|---|---|---|---|---|
| Theoriespezifische Erfassung epistemologischer Überzeugungen | Trautwein & Lüdke (2004) | n.a. | Dualismus, Relativismus (Theoriespezifische epistemologische Überzeugungen und Grenzen wissenschaftlicher Erkenntnis: Einschätzung von 10 Theorien aus vier Fachgebieten mittels je 9 Items (Biologie: Selektionstheorie, Theorie zum Aussterben der Dinosaurier, Theorie zur Abstammung der Menschen aus Afrika; Medizin: Cholesterin-Theorie, Theorie zur karzinogenen Wirkung des Rauchens, Prionentheorie der BSE-Entstehung; Physik: Urknall-Theorie, Relativitätstheorie; Psychologie: Modelllernen bei Gewaltcomputerspielen, Wirkung von Hormonen auf das räumliche Denkvermögen) BSE-Theorie (9, .78), Urknalltheorie (9, .78), restliche Theorien (durchschnittlich .84) (Varianzaufklärung: 29%) | 9 | n.a. | theoriespezifisch zu Disziplinen Psychologie, Physik, Biologie, Medizin | Studierende |
| deutsche Version des DEBQ | Richter (2004) | Hofer (2000) | Sichertheit/Einfachheit, Begründung: Persönlich, Quelle: Autorität, Erreichbarkeit der Wahrheit | 18 | 5-stufige Ratingskala | anpassbar an Studienfach/ an Teilbereich aus Disziplin | Studierende |
| Studienfachspezifische epistemologische Überzeugungen | Trautwein & Lüdke (2004) | Hofer (2000) | Einstellungen ggü. Veränderlichkeit der Erkenntnismöglichkeiten im Studienfach (3, .74) | 3 | n.a. | anpassbar an Studienfach ("Denken Sie nun bitte ganz speziell an Ihr [Haupt-]Studienfach") | Studierende |
| Connotative Aspects of Epistemological Beliefs (CAEB) | Bartholome et al. (2006); Stahl & Bromme (2007) | SEQ, Schommer (1990); DEBQ, Hofer (2000); EBI, Jacobson & Jehng (1999); EBAPS, Elby et al. (2003) | Textur (10, .78-.86), Variabilität (7, .69-.79) (Varianzaufklärung: 49%, 62%) | 17 Adjektivpaare (Original 24) | 9-stufige Ratingskala, semantisches Differential | Naturwissenschaften: Biologie (Pflanzenkunde, Genetik), organische Chemie, Physik | Studierende |
| Fragebogen zur Erfassung des Entwicklungsniveaus epistemologischer Überzeugungen (FREE) | Krettenauer, 2005) | Boyes & Chandler (1992), King & Kitchener (1994) | Absolutismus, Relativismus, Postrelativismus | 3 Aussagen mit Konflikten auf unterschiedlichen Niveaus | 6-stufige Ratingskala, halbstandardisiertes Interview | Naturwissenschaften, Sozialwissenschaften, Geisteswissenschaften | Schülerinnen und Schüler, Studierende |
| Übersetzung der Certain Epistemological Belief Scale von Conley et al. (2004) | Urhahne & Hopf (2004); Urhahne (2006) | Conley et al. (2004); Elder (2002); Hofer & Pintrich (1997) | Sicherheit des Wissens (6, .41 - .76), Wissensquelle (5, .67-.72), Rechtfertigung des Wissens (9, .61-.68), Entwicklung des Wissens (6, .66-.69) (Varianzaufklärung: 39,8%) | 26 | 5-stufige Ratingskala | Naturwissenschaft | Schülerinnen und Schüler, Studierende |
| deutsche Übersetzungen des DEBQ | Pieschl, Stahl & Bromme (2006) | DEBQ von Hofer (2000); Hofer & Pintrich (1997) | resultierende DEBQ Dimensionen: Sicherheit (9, .74), Zugänglichkeit (6, .58) (Varianzaufklärung: 36%) | 15 | n.a. | themenspezifisch anwendbar | Studierende |

ABBILDUNG 2.16: Bedeutende deutschsprachige Fragebögen zur mehrdimensionalen, disziplinspezifischen Erfassung epistemologischer Überzeugungen

Für die Darstellung weiterer Instrumente sei auf Buehl (2008), Priemer (2006), Zinn (2013) und Mokwinski (2011) verwiesen. Buehl (2008, S. 70 ff) bietet eine umfassende, tabellarische Übersicht über die Studien, in deren Rahmen die wichtigsten Fragebögen zur Erfassung epistemologischer Überzeugungen entstanden sind. Hier werden außerdem die theoretische Fundierung der Instrumententwicklungen, die resultierenden Dimensionen sowie einige zentrale psychometrische Eigenschaften und zentrale Informationen der resultierenden Fragebögen berichtet. Für eine ergänzte Zusammenstellung der ausschließlich deutschsprachigen Fragebogenverfahren zur Erfassung epistemologischer Überzeugungen aus der Basis von Priemer (2006, S. 169 f.) sei auf Zinn (2013, S. 310 f.) verwiesen. Für eine Übersicht ausschließlich disziplinspezifischer Fragebögen bietet sich die Übersicht bei Mokwinski (2011, S. 39) an. Eine tabellarische Übersicht zu Studienergebnissen mit Einsatz des EBI ist ebenfalls bei Mokwinski (2011, S. 37) zu finden.

Wie aus den vorherigen Darstellungen hervorgeht, existieren zahlreiche quantitative Erhebungsinstrumente. Der Einsatz quantitativer Verfahren ermöglicht einerseits die Hypothesenprüfung und erlaubt es, auf Basis der erhobenen Daten Rückschlüsse auf die Population zu ziehen (Döring & Bortz, 2016). Auch sind standardisierte schriftliche Befragungen forschungsökonomischer. Die Daten größerer Stichproben können in kürzerer Zeit erhoben und ausgewertet werden (Mason, 2016). Der Einsatz von Fragebögen ermöglicht zudem Verbindungen zu anderen psychologischen Konstrukten wie Motivation, Fachinteresse, Lernstrategien und fachspezifisches Selbstkonzept einfacher zu erforschen (ebd.).

### 2.9.3 Probleme der bestehenden quantitativen Erhebungsinstrumente

Eine Problematik der bisherigen quantitativen Erhebung epistemologischer Überzeugungen geht auf die theoretische Konzeptualisierung zurück. Der Großteil der existierenden Fragebögen geht auf Schommers (1990) Konzeption zurück und beinhaltet demnach nicht nur Überzeugungen zum Wissen und Wissensprozess, sondern auch Überzeugungen über das Lernen und Intelligenz, die nicht zum Kernkonstrukt epistemologischer Überzeugungen gehören (Bråten & Olaussen, 2005; Hofer & Pintrich, 1997). Schommer-Aikins selbst betrachtet die Überzeugungen zu Lernen und Intelligenz in der Weiterentwicklung ihrer Konzeption (Schommer-Aikins, 2004) als explizit von den epistemologischen Überzeugungen getrennt. Wo sie vorher von vier Dimensionen epistemologischer Überzeugungen sprach, differenziert sie nun Überzeugungen zu Wissen und Überzeugungen zu Lernen. Schommer (1990) fand eine vierfaktorielle Stuktur und konnte die Dimension *Quelle des Wissens*, die unumstritten zum Kernbereich der epistemologischen Überzeugungen gehört (Hofer & Pintrich, 1997), nicht abbilden. In den nachfolgenden Studien fand Schommer teilweise nur eine dreifaktorielle Struktur und später wieder die vierfaktorielle Struktur (Schommer et al., 1992). Zwar liegen Studien vor, die beim Einsatz des SEQ die gleichen vier von Schommer (1990) gefundenen Faktoren (*simple knowledge, certain knowledge, innate ability* und *quick learning*) replizieren konnten (z.B. Bendixen et al., 1994), allerdings gibt es zahlreiche Studien, in denen dies nicht gelang und in denen inhaltlich andere Faktoren identifiziert wurden, wie *depend on authority* (Kardash & Scholes, 1996), *dependence on authority* und *attainability of truth* (Mori, 1999), *avoid integration*, (Kardash & Noel, 2000), *fixed/innate ability* als kombinierter Faktor

(Chan & Elliott, 2000) sowie *knowledge construction and modification* und *control of knowledge acquisition* (Bråten & Olaussen, 2005). Insgesamt zeigen diese Ergebnisse, dass die Faktorstruktur nicht sehr stabil ist. Immer wieder zeigen sich beim Einsatz des SEQ Faktoren, die Überzeugungen zur Quelle des Wissens abdecken, wie beispielsweise der Faktor *dependence on authority* von Mori (1999) oder *depend on authority* (Kardash & Scholes, 1996). Das erklärt auch, wieso auf Basis des SEQ so viele eigene Instumententwicklungen vorgenommen wurden (z.B. Chan & Elliott, 2002; Cole & Goetz, 2000; Jehng et al., 1993; Schraw et al., 1995), statt den SEQ weiterhin einzusetzen und weshalb sich der EBI so großer Beliebtheit erfreut, weiterentwickelt und immer häufiger eingesetzt wurde.

Eine weitere grundlegende Problematik betrifft den SEQ, JEQ, EBI und generell alle Erhebungsinstrumente, die in den Abbildungen 2.13 (S. 116) und 2.15 (S. 120) dargestellt sind. All diese Fragebögen erfassen epistemologische Überzeugungen disziplinübergreifend; und zwar unabhängig vom Forschungsinteresse, d.h. auch wenn die Erfassung auf die epistemologischen Überzeugungen von Lernenden in einer spezifischen Disziplinen abzielt. Mittlerweile herrscht jedoch Konsenz, dass epistemologische Überzeugungen sowohl disziplinspezifisch als auch disziplinübergreifend erfasst werden können (Buehl & Alexander, 2001; Buehl et al., 2002; Hofer, 2000; Hofer & Pintrich, 1997; Pintrich, 2002; Schommer-Aikins & Hutter, 2002), da Lernende sowohl über allgemeine als auch über disziplinspezifische Überzeugungen zu Wissen und dem Wisensprozess verfügen (Muis et al., 2006). Die undifferenzierte Anwendung des SEQ und verwandter Instrumente, das heißt von Fragebögen, deren Items allgemein gehalten sind und keinen Bezug zur Disziplin herstellen, mag auch zur Erklärung der instabilen Faktorenstruktur und eingeschränkten Reliabilität beitragen (vgl. z.B. Clarebout et al., 2001; Stahl et al., 2006). Dies führt dann auch dazu, dass teilweise ein großer Teil der Items von den Analysen der Auswertung ausgeschlossen werden muss (vgl. Bråten & Strømsø, 2004; Trautwein et al., 2004). Für diese mangelnde Messzuverlässigkeit werden die Unschärfe und die interpretative Weite der allgemeinen epistemologischen Überzeugungen verantwortlich gemacht (vgl. Bromme & Stahl, 2005; Trautwein & Lüdtke, 2004, vgl. Zinn, 2013, S. 116).

Eine weitere Problematik, die sowohl die disziplinübergreifenden als auch disziplinspezifischen Fragebögen betrifft, steht in Verbindung mit der Kontextabhängigkeit epistemologischer Überzeugungen (vgl. Bråten & Strømsø, 2010; Haerle & Bendixen, 2008; Muis, 2008; Muis et al., 2006 und siehe Kapitel 2.5 und 2.6). Die prominentesten Fragebögen – der SEQ, JEQ, EBI, DEBQ und DSBQ – wurden alle in den USA entwickelt und validiert. Sie wurden in einem spezifischen akademischen und kulturellen Kontext mit einer bestimmten Zielgruppe entwickelt und eingesetzt, aber vielfach mittels einer einfachen Übersetzung der Items in andere kulturelle Kontexte, wie Belgien, Norwegen, Süd Korea und China (z.B. Bråten & Olaussen, 2005; Chan & Elliott, 2000; Clarebout et al., 2001; Youn, 2000), übertragen (vgl. Hofer, 2008). Die Erhebungsinstrumente sind allerdings nicht notwendigerweise auf andere kulturelle Kontexte und Zielgruppen übertragbar (vgl. Chan & Elliott, 2002; Hofer, 2001). Das geht schon aus der theoretischen Konzeption des TIDE Frameworks (Muis et al., 2006) hervor. Es beschreibt unter anderem, dass die Entwicklung individueller epistemologischer Überzeugungen vom soziokulturellen Kontexts beeinflusst wird und dass auch die

situative Aktivierung spezifischer wissensbezogener Überzeugungen kontextabhängig ist (siehe Kap. 2.6, S. 75). In verschiedenen Studien konnten bereits Hinweise auf den Einfluss der Kultur auf die epistemologischen Überzeugungen gefunden werden und es ergaben sich beim Einsatz der Fragebögen in anderen kulturellen Kontexten (z.B. Italien, Spanien, Süd Korea, Oman, Norwegen) teilweise andere Faktorstrukturen, stark abweichende Reliabilitäten und Varianzaufklärungen (Bråten & Olaussen, 2005; Cano, 2005; Karabenick, 2005; Mason, 2000; Qian & Alvermann, 1995; Youn, 2000).

Der DEBQ von Hofer (2000) hat den Vorteil, nur die Kerndimensionen epistemologischer Überzeugungen abzubilden und keine Überzeugungen zu Lernen zu beinhalten. Der DEBQ wurde anhand einer Stichprobe von Psychologiestudierenden entwickelt, die zu ihren wissensbezogenen Überzeugungen zu Psychologie und Naturwissenschaft befragt wurden. Er erhebt aber den Anspruch, auch für andere Disziplinen einsetzbar zu sein, indem die Befragten im Rahmen der Instruktion darauf hingewiesen werden, auf welche Disziplin sie die Items beziehen sollen. Die Items enthalten wiederum die Formulierung „in this field“. Es ist allerdings fraglich, ob diese Anpassung tatsächlich funktioniert. Der bloße Hinweis zu Beginn einer Erhebung, an ein Fach zu denken, mag zu schwach sein, um es mit den starken epistemologischen Überzeugungen zum akademischen Schwerpunktbereich einer Person aufnehmen zu können, die zudem mit Fortschreiten des Studiums über die akademische Sozialisation im Studien hinweg noch immer stärker werden (vgl. TIDE Modell, Muis et al., 2006). Auch die Problematik der fraglichen Übertragbarkeit in einen anderen kulturellen Kontext betrifft den DEBQ.

Bei Fragebögen, denen die Annahme zugrundeliegt, sie seien vereinbar mit der entwicklungspsychologischen Sichtweise auf epistemologische Überzeugungen, stellt sich die Frage nach der inhaltlichen Interpretierbarkeit der Selbstauskünfte mittels Likertskalen (Mason, 2016). Hofer (2000) beispielsweise sieht ihren DEBQ auf Itemebene als vereinbar mit der Annahme, dass sich epistemologische Überzeugungen über die Zeit von absolutistischen, über multiplizistische hin zu evaluatistischen Überzeugungen entwickeln. Wird aber nun davon ausgegangen, dass der linke Pol eines likertskalierten Items der absolutistischen Überzeugung entspricht und der rechte Pol eine evaluatistische, d.h. sophistiziertere Überzeugung widerspiegelt, ergeben sich Probleme mit der inhaltlichen Interpretierbarkeit der mittleren Ausprägung (Mason, 2016). Dieser Kritikpunkt betrifft also die Validität der Erhebung in dem Sinne, dass fraglich wird, ob das Item das auch wirklich misst, was gemessen werden soll. Wenn beispielsweise der eine Pol der Überzeugung entspricht, dass Wissen von externen Autoritäten übertragen wird und somit dem Absolutismus, während der andere Pol der Überzeugung entspricht, dass Wissen selbst in Interaktion mit anderen konstruiert wird und somit dem Evaluatismus, über welche Überzeugung verfügt dann eine Person, die genau die Mitte ankreuzt (Muis et al., 2006)? Ginge man von einer linearen Entwicklung der epistemologischen Überzeugungen aus, welche die Likertskalierung ja impliziert, müsste diese Mitte der Entwicklungsstufe des Relativismus entsprechen. Es ist jedoch fraglich, ob diese Interpretation zulässig ist. Kritisiert wurde nicht nur die Interpretierbarkeit der mittleren Ausprägung, sondern es wurde auch angezweifelt, ob nicht der rechte Pol generell

eher einer relativistischen Sichweise auf Wissen entspricht und der linke Pol einer absolutistischen (Hofer, 2016). Somit würde der rechte Pol zwar immer noch die sophistiziertere Überzeugung abbilden, allerdings wäre die höchste Entwicklungsstufe, der Evaluatismus, gar nicht mehr abgebildet. Betrachtet man beispielsweise die Überzeugungen zur Sicherheit des Wissens, würde die Annahme, dass Wissen komplett sicher ist, dem wenig sophistizierten Absolutisten entsprechen. Die gegenteilige Annahme, dass Wissen komplett unsicher ist, bildet allerdings keine evaluatistische sondern höchstes eine relativistische Stufe der epistemologischen Entwicklung ab. Somit würde ein entsprechendes Item nur zwei der drei Entwicklungsschritte abbilden (ebd.).

Wie im vorangegangenen Kapitel dargestellt wurde, existieren einige quantitative Erhebungsinstrumente zur Erfassung epistemologischer Überzeugungen. Viele der bestehenden Fragebögen sind jedoch disziplinübergreifend und werden nach wie vor auch in Kontexten eingesetzt, in denen das zugrundeliegende Forschungsinteresse vielmehr nach der Nutzung eines disziplinspezifischen Erhebungsinstruments verlangen würde. Die existierenden disziplinspezifischen Erhebungsinstrumente beziehen sich zum Großteil auf die naturwissenschaftliche Disziplin. Fast alle der bestehenden disziplinübergreifenden und -spezifischen Instrumente beinhalten Skalen zur Erfassung der Überzeugungen zum Lernen, die nicht zum Kernkonstrukt epistemologischer Überzeugungen gehören. Dies trifft auch auf die deutschen Fragebögen zur Erfassung epistemologischer Überzeugungen zu.

## 2.10 Synthese

In den vorangegangenen Kapiteln wurde sowohl die Forschungsentwicklung dargestellt, als auch anhand der verschiedenen Konzeptualisierungen aufgezeigt, was unter epistemologischen Überzeugungen jeweils verstanden wird. Es wurden zentrale Forschungsdiskurse wiedergegeben und die Befundlage dargestellt. Es wurde außerdem gezeigt, wie epistemologische Überzeugungen entstehen, sich entwickeln, verändern, sich in Theorien der Kognition und Metakognition verorten lassen und welche Verbindungen sie mit dem Lehr- und Lernprozess ausweisen. Diese vorgenommene Eingrenzung des Zielkonstrukts stellt einen wichtigen Arbeitsschritt der Instrumententwicklung dar, der zur Erstellung des Testentwurfs nötig ist (vgl. Bühner, 2011, S. 97 ff.). Auf Basis dieses theoretischen Hintergrundes wird im Folgenden präzisiert, wie das Konstrukt epistemologischer Überzeugungen im Rahmen dieser Arbeit positioniert wird, d.h. der Messgegenstand der Instrumententwicklung wird hier definiert (vgl. Bühner, 2011, S. 105 ff.). Im Anschluss wird die besondere Bedeutung epistemologischer Überzeugungen für Lehramtsstudierende aufgezeigt und zu den Fragestellungen dieser Arbeit übergeleitet, indem die Relevanz der Entwicklung eines neuen Erhebungsinstrumentes zur Erfassung der Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten hergeleitet wird.

Den theoretischen Kern dieser Arbeit bilden:

1. Die Konzeptualisierung des *Epistemological Belief System* von Schommer (1990), auf welche die explizite Mehrdimensionalität zurückgeht,
2. die Kerndimensionen epistemologischer Überzeugungen unter Ausschluss der Überzeugungen zum Lernen nach der Konzeptualisierung der *Epistemological Theories* von (Hofer & Pintrich, 1997), unter Berücksichtigung von Conley et al. (2004),
3. das Modell von Hofer (2004a) zur Verortung der epistemologischen Überzeugungen in Theorien der Kognition und Metakognition,
4. das *Integrative Model for Personal Epistemology* (IM) von Bendixen und Rule (2004) und Rule und Bendixen (2010) zur Erklärung des Veränderungsmechanismus epistemologischer Überzeugungen,
5. die Kontextabhängigkeit, die Disziplinspezifität sowie die Annahmen zur Entstehung epistemologischer Überzeugungen nach der *Theory of Integrated Domains in Epistemology* (TIDE) von Muis et al. (2006).
6. Das Modell von Hofer (2001) und das *Embedded Model* Schommer-Aikins (2004) dienen zur Erklärung der Beziehung zwischen epistemologischen Überzeugungen und dem Lehr-Lernprozess.

Das Konstrukt epistemologischer Überzeugungen besteht, zurückgehend auf Schommer (1990), aus mehreren Dimensionen. Dieses System von Überzeugungen ist nach Hofer und Pintrich (1997) theorieähnlich organisiert und betrifft übergeordnet Überzeugungen zu Wissen und dem Wissensprozess, während die Überzeugungen zum Lernen nicht in den Kernbereich epistemologischer Überzeugungen fallen (vgl. Hofer, 2005). Die Kerndimensionen sind Überzeugungen zur *Quelle des Wissens*, *Rechtfertigung des Wissen*, *Sicherheit des Wissens* und *Einfachheit des Wissens* nach Hofer und Pintrich (1997). Es gibt Hinweise darauf, dass sich die Dimension *Entwicklung des Wissens* nach Conley et al. (2004) als eigene Dimension, getrennt von den übrigen vier Dimensionen, identifizieren lässt. Die Überzeugungen zur Stabilität, Struktur und Entwicklung des Wissens können dem übergeordneten Bereich *Beschaffenheit des Wissens* zugeordnet werden, während die Überzeugungen zur Quelle und Rechtfertigung des Wissens zum Bereich *Beschaffenheit des Wissensprozesses* gehören.

Individuen verfügen sowohl über:

1. allgemeine, übergreifende nicht-akademische Überzeugungen (Alltag, Arbeitsumfeld etc.),
2. allgemeine akademische Überzeugungen als auch über
3. disziplinspezifische akademische Überzeugungen (Muis et al., 2006).

Die wissensbezogenen Überzeugungen einer Person werden im instruktionalen, akademischen und soziokulturellen Kontext aktiviert und von diesem beeinflusst (Hofer, 2005, 2006b; Muis et al., 2006). Sie entstehen über die gesamte Lebenspanne (Muis et al., 2006) in dynamischer Interaktion des Individuums mit dem alltäglichen sozialen, kulturellen und akademischen Umfeld (Baxter Magolda, 2004; Belenky et al., 1986; Bendixen & Rule, 2004; Hofer & Pintrich, 1997) und werden von Erziehung, Bildung, Kultur sowie akademischen Erfahrungen beeinflusst (King & Kitchener, 1994; Kuhn, 1991; Kuhn & Pearsall, 2000; Muis et al., 2006), was zu einer Weiterentwicklung der allgemeinen nicht-akademischen, der allgemeinen akademischen sowie der disziplinspezifischen epistemologischen Überzeugungen einer Person führt (IM, Bendixen & Rule, 2004, TIDE, Muis et al., 2006 sowie Hofer, 2004a). Mit Eintritt in das Bildungssystem und mit fortschreitenden Bildungserfahrungen entwickelt ein Individuum in Auseinandersetzung mit verschiedenen Unterrichtsfächern (in der Schule) und akademischen Disziplinen (in der Universität) parallel jeweils verschiedene disziplinspezifische epistemologische Überzeugungen (TIDE, Muis et al., 2006). Mit zunehmender akademischer Erfahrung und Spezialisierung in einer bestimmten Disziplin nimmt der Einfluss dieser disziplinspezifischen wissensbezogenen Überzeugungen auf die allgemeinen akademischen Überzeugungen zu (TIDE, Muis et al., 2006). Die Entwicklung der drei Überzeugungssysteme (allgemein nicht-akademisch, allgemein akademisch, disziplinspezifisch akademisch), die jeweils Überzeugungen zu allen Dimensionen enthalten, verläuft von weniger sophistizierten bis hin zu sophistizierten individuellen wissensbezogenen Überzeugungen (IM, Bendixen & Rule, 2004; Rule & Bendixen, 2010 sowie auch Hofer, 2005; King & Kitchener, 1994; Kuhn, 1991; Kuhn & Pearsall, 2000; Perry, 1970), wobei rückläufige Entwicklungen auch auftreten und verursacht werden können (Bendixen & Rule, 2004; Muis et

al., 2006; Rule & Bendixen, 2010). Jede der Überzeugungsdimensionen entwickelt sich von einer weniger sophistizierten bis hin zu einer sophistizierteren Perspektive auf Wissen und den Wissensprozess (Bendixen & Rule, 2004; Muis et al., 2006; Rule & Bendixen, 2010). In der weniger sophistizierten Perspektive wird Wissen als feststehend, unveränderlich, aus einfachen, unverbundenen Fakten bestehend, von Autoritäten vermittelt und als zweifelsfrei wahr oder falsch klassifizierbar betrachtet, ohne einer kritischen Auseinandersetzung mit den Wissensquellen zu bedürfen (Bendixen & Rule, 2004; Hofer & Pintrich, 1997; Muis et al., 2006). In der sophistizierteren Ausprägung wird Wissen als vorläufig, veränderlich, komplex miteinander verbunden, persönlich in Interaktion mit dem sozialen Umfeld konstruiert und kritisch zu hinterfragen und zu rechtfertigen betrachtet (Bendixen & Rule, 2004; Hofer & Pintrich, 1997; Muis et al., 2006). Die individuellen epistemologischen Überzeugungen sind Teil des metakognitiven Bewusstseins und werden bei allen Lernprozessen (Hofer, 2004a) und zwar in Abhängigkeit vom jeweiligen soziokulturellen, akademischen und instruktionalen Kontext aktiviert (Hofer, 2005, 2006b; Muis et al., 2006). Sie operieren sowohl auf kognitiver als auch metakognitiver Ebene (Bendixen & Rule, 2004; Hofer, 2004a, 2005, 2016; Muis et al., 2006), da die Metakognition auf die Kognition zugreift und diese wiederum speist (Veenman et al., 2006).

Ein sophistizierteres epistemologisches Überzeugungssystem wirkt sich dabei förderlich auf multiple lernrelevante Faktoren und demnach förderlich auf den Lehrprozess eines Individuums aus, während weniger fortgeschrittene Überzeugungen eher lernhemmend sind (Hofer, 2001; Schommer-Aikins, 2002). Sophistizierte epistemologische Überzeugungen sind lernförderlich im Hinblick auf multiple lernrelevante Faktoren, wie Lernemotionen (z.B. Muis et al., 2015), intrinsische Motivation (z.B. Buehl & Alexander, 2005; Hofer, 1999), Selbstkonzept (z.B. Urhahne & Harms, 2006), Fachinteresse, Selbststeuerungsfähigkeit (z.B. Bromme, 2010), Kontroll- und Selbstwirksamkeitsüberzeugungen (z.B. Muis & Foy, 2007; Paulsen & Feldman, 1999) und unterstützen den Lerntransfer (Jacobson & Spiro, 1995), die Informationssuche (z.B. Mason et al., 2010b; Pieschl et al., 2013), das Textverstehen (z.B. Bråten & Strømsø, 2004, 2006; Bråten et al., 2008; Schommer-Aikins, 2004) und den *Conceptual Change* (Franco et al., 2012). Sie führen insbesondere zu förderlichem Lernverhalten, indem passendere Lernstrategien eingesetzt werden (Hofer, 2001; Schommer-Aikins, 2002).

Die übergeordnete Zusammenhangsstruktur zwischen sophistizierten und wenig sophistizierten epistemologischen Überzeugungen und Lernstrategien, die sich aus den oben dargestellten Forschungsbefunden (siehe Abschnitt 2.8.2, ab S. 96) ergibt, wird in Abbildung 2.17 (S. 129) zusammengefasst. Hieraus geht hervor, dass sophistizierte epistemologische Überzeugungen mit dem Einsatz tiefenorientierter Lernstrategien einhergehen, aber negativ mit der oberflächenorientierten Wiederholungsstrategie zusammenhängen (siehe Abb. 2.17, S. 129). Genau umgekehrt verhält es sich mit den wenig sophistizierten epistemologischen Überzeugungen, die negativ mit tiefenorientierten Lernstrategien korrelieren, aber positiv mit der oberflächenorientierten Wiederholungsstrategie (siehe Abb. 2.17, S. 129).

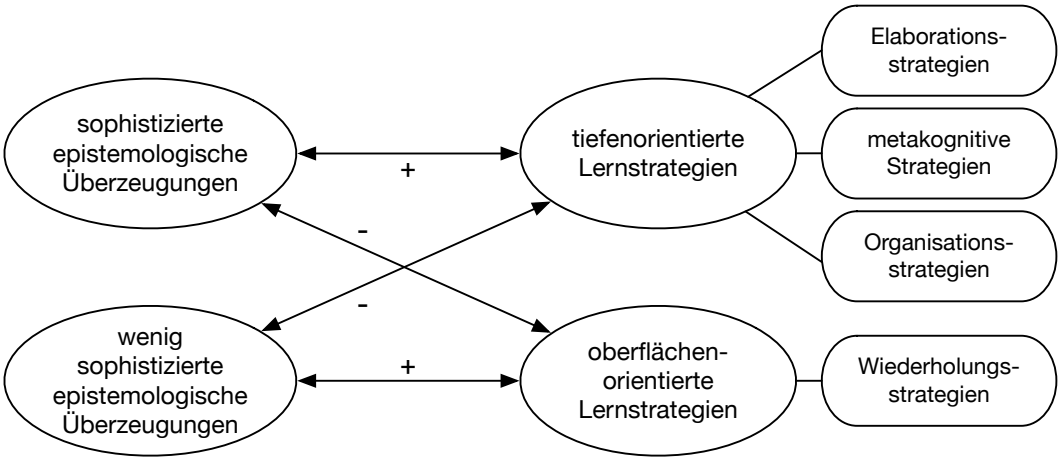


ABBILDUNG 2.17: Zusammenhangsstruktur zwischen sophistizierten und wenig sophistizierten epistemologischen Überzeugungen und Lernstrategien

Die aus den Forschungsbefunden (siehe Abschnitt 2.8.2, ab S. 96) abgeleiteten Zusammenhänge zwischen den Dimensionen epistemologischer Überzeugungen in sophistizierter Ausprägung und den verschiedenen Lernstrategiedimensionen sind in Abbildung 2.18 (S. 130) dargestellt. Neben der übergeordneten Zusammenhangsstruktur lassen sich die Verbindungen zwischen den Dimensionen epistemologischer Überzeugungen und Lernstrategien noch genauer betrachten. Die grünen Pfeile symbolisieren dabei positive Zusammenhänge, während die roten für negative Zusammenhänge stehen. Die dargestellten Zusammenhänge basieren auf den oben dargestellten, vielfältigen Forschungsergebnissen. Die negative Verbindung der Dimensionen *Überzeugungen zur Quelle des Wissens* und *Entwicklung des Wissens* mit den *Wiederholungsstrategien* ist allerdings weniger stark fundiert, da dazu weniger Forschungsergebnisse vorliegen.

Gerade für Studierende, die im Rahmen ihres Studiums beständig selbstgesteuert lernen müssen (Schworm & Gruber, 2017), sind die epistemologischen Überzeugungen lernrelevant, da sich ein sophistiziertes epistemologisches Überzeugungssystem positiv auf alle Phasen des selbstregulierten Lernens auswirkt (z.B. Bromme, 2010; Bromme et al., 2014; Stahl & Bromme, 2007). Wissensbezogene Überzeugungen wirken als Linse, durch die Lernende die Lernaufgabe wahrnehmen sowie Lerninhalte erfassen und dienen der erfolgreichen Anpassung des Lernverhaltens – im Sinne einer förderlichen Lernstrategieauswahl – an die wahrgenommenen Anforderungen und Rahmenbedigungen der Aufgabe (Bromme, 2010; Pieschl et al., 2013).

Für Lehramtsstudierende sind die eigenen epistemologischen Überzeugungen in zweifacher Hinsicht von besonderer Bedeutung, insofern sie sowohl deren eigenes Lernen im Studium als auch deren Konzeption von Lernkontexten für das Lernen anderer beeinflussen (Buehl & Fives, 2016, S. 248). Die Lehrerbildung erstreckt sich über einen langen Zeitraum und umfasst in Deutschland das Studium, Praktika und das Referendariat. Angehende Lehrkräfte müssen also kontinuierlich selbstgesteuert lernen – und zwar sowohl in

den verschiedenen Wissensdomänen ihrer Unterrichtsfächer als auch im erziehungswissenschaftlichen Wissensbereich – um ihr professionelles Handeln weiterzuentwickeln und das Lernen von Schülerinnen und Schülern erfolgreich fördern zu können (Buehl & Fives, 2016).

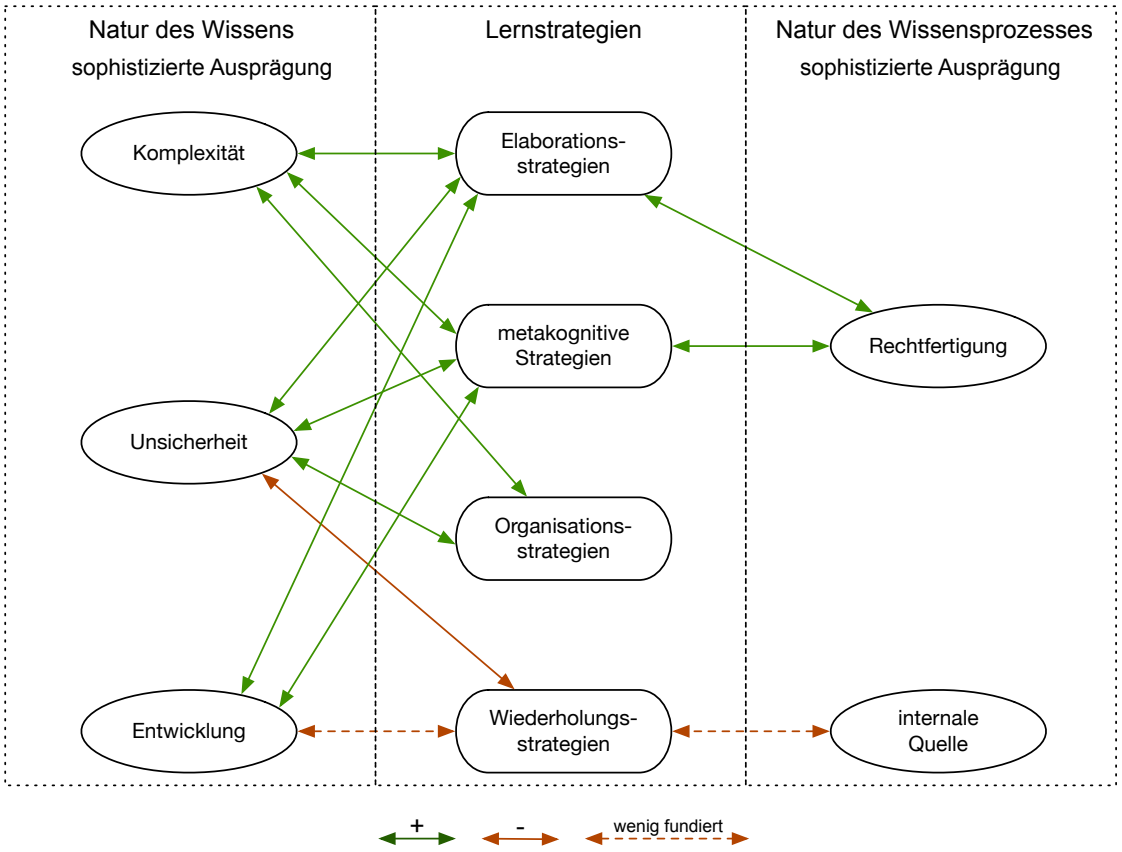


ABBILDUNG 2.18: Zusammenhangsstruktur zwischen den sophistizierten Ausprägungen epistemologischer Überzeugungen und Lernstrategien

Aber auch als praktizierende Lehrkräfte wird implizit im Berufsalltag und explizit in Fort- und Weiterbildungen beständig weitergelernt. Die wissensbezogenen Überzeugungen filtern, wie Lehramtsstudierende das Lernen in ihren universitären Veranstaltungen erleben und inwiefern sie sinnvolle, tiefenorientierte Lernansätze betreiben (Many et al., 2002; Muis, 2004; Peng & Fitzgerald, 2006; Yadav & Koehler, 2007). Sie wirken sich folglich darauf aus, wie tief ihr in den universitären Veranstaltungen (*teacher education courses*) erworbenes Verständnis geht und beeinflussen außerdem ihre Entscheidungsfindung, Planung, Durchführung und Bewertung der praktischen Anwendung des Gelernten (Buehl & Fives, 2016; Yadav & Koehler, 2007). Schon bei angehenden Lehrerinnen und Lehrern stehen die individuellen wissensbezogenen Überzeugungen in starker Verbindung mit der unterrichtlichen Zielsetzung und den im Unterricht umgesetzten Handlungen (z.B. im Rahmen von Praktika oder des Referendariats) (z.B. Kang, 2008; Yadav & Koehler, 2007) und sophistizierte epistemologische Überzeugungen sind mit einer stärker konstruktivistisch ausgeprägten Didaktik verbunden (Brownlee et al., 2016). Angehende Lehrkräfte mit sophistizierteren epistemologischen Überzeugungen nutzen eher pädagogische Praktiken, die die Kinder ermutigen,

unabhängig zu denken und Wissen zu konstruieren als Lehrkräfte mit weniger sophistizierten epistemologischen Überzeugungen (Brownlee et al., 2016).

Auch die selbstständige Recherche und das selbstständige Lesen und Erarbeiten von Texten, die wichtige Aspekte eines Studiums darstellen, sind vom epistemologischen Überzeugungssystem abhängig. Sophistizierte epistemologische Überzeugungen beeinflussen das Verhalten von Lernenden bei der selbstregulierten Informationssuche und -bewertung positiv (Hofer, 2004b; Mason et al., 2010b; Pieschl et al., 2013; Whitmire, 2004). Je fortgeschrittener das epistemologische Überzeugungssystem ist, desto eher wenden Studierende viel Zeit und Sorgfalt bei der Planung der Informationssuche auf, nutzen multiple Quellen, differenzieren zwischen der Herkunft von Inhalten und vergleichen kritisch zwischen verschiedenen Inhalten. Sophistizierte epistemologische Überzeugungen wirken sich förderlich auf die Textverarbeitung und das Textverstehen aus (z.B. Bråten & Strømsø, 2004; Kardash & Noel, 2000; Schommer et al., 1992; Schommer & Walker, 1995; Schommer-Aikins, 2004; Schraw et al., 2002). Für Lehramtsstudierende ist es folglich von großer Bedeutung, über sophistizierte epistemologische Überzeugungen zu verfügen,

> „that can influence the depth of understanding they take from teacher education courses, the development of deep and varied knowledge systems for teaching, and the use of effective, informed decision-making across a range of contexts“ (Brownlee et al., 2011, S. 303).

Da epistemologische Überzeugungen aber disziplinspezifisch sind (Muis et al., 2006), wäre es erstrebenswert, dass Lehramtsstudierende sowohl bezüglich der Wissensdomänen ihrer jeweiligen Unterrichtsfächer als auch bezüglich des erziehungswissenschaftlichen Wissens über sophistizierte Überzeugungsysteme verfügen.

Aus erstens der dargestellten Bedeutung epistemologischer Überzeugungen für den Lehr-Lernprozess, zweitens der lernförderlichen Natur sophistizierter epistemologischer Überzeugungen sowie drittens der besonderen Relevanz der wissensbezogenen Überzeugugen von angehenden Lehrkräften, lässt sich folgern, dass die Förderung sophistizierter epistemologischer Überzeugungen – und zwar sowohl für die Fachdisziplin(en) als auch die Erziehungswissenschaft – in der Lehrerbildung ein zentrales Ziel darstellen sollte (vgl. Brownlee et al., 2011, 2016). Es ist außerdem anzunehmen, dass die erfolgreiche Förderung sophistizierter epistemologischer Überzeugungen bei Lehramtsstudierenden in der Lehrerbildung sowohl die Qualität der Lehrerbildung selbst verbessern würde (Brownlee et al., 2011, S. 303) als auch die Schulentwicklung sowie Schul- und Bildungsreformen fördern könnte (Chai et al., 2006). Denn Lehrkräfte mit elaborierteren Überzeugungssystemen akzeptieren Schulreformen, die eine Anpassung des pädagogischen Handelns erfordern, eher, als Lehrkräfte mit weniger fortgeschrittenen epistemologischen Überzeugungen (Patrick & Pintrich, 2001). Für eine Förderung schon während des Studiums spricht zusätzlich, dass Lehramtstudierende ihr epistemologisches Überzeugungssystem eher in Richtung einer konstruktivistischen didaktischen Sichtweise ändern, als bereits praktizierende Lehrkräfte (Tsai, 2006).

Es ist für die Lehrerbildung zunächst wichtig, das vorherrschende epistemologische Überzeugungssystem von Lehramtsstudierenden zu verstehen, um sophistizierte epistemologische Überzeugungen zielgerichtet fördern zu können (Chai et al., 2006). Eine Voraussetzung der zielgerichteten Förderung ist also die Identifikation des IST-Zustandes. Nur wenn das aktuell vorherrschende epistemologische Überzeugungssystem ermittelt werden kann, lässt sich feststellen, bei wem und in Bezug auf welche Überzeugungsdimensionen Förderbedarf vorliegt. Und nur wenn die epistemologischen Überzeugungen erhoben werden können, kann der Erfolg einer durchgeführten Förderungsmaßnahme ermittelt werden. Hier eignen sich quantitative Erhebungssinstrumente besser als qualitative, da sie ökonomisch einsetzbar und auszuwerten sind, sich gut für die Erfassung von Veränderungen in Merkmalsausprägungen eignen sowie die einfache Untersuchung von Verbindungen zu anderen lernrelevanten Konstrukten ermöglichen (Mason, 2016).

Für die verschiedenen Fachdisziplinen, wie beispielsweise Mathematik, Physik und Geschichte ist die quantitative Erhebung mittels Fragebogen bereits möglich, da es passende Instrumente gibt. Es gibt jedoch bisher kein quantitatives Erhebungsinstrument, das sich explizit für die Erfassung der Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen und Wissensprozess eignet. Und das, obwohl die Erfassung dieser epistemologischen Überzeugungen von großer theoretischer und praktischer Relevanz für angehende Lehrerinnen und Lehrer und auch für deren Schülerinnen und Schüler ist. Dieses erziehungswissenschaftliche Wissen umfasst schließlich pädagogische und didaktische Inhalte, die für den späteren Lehrberuf zentral sind, wie z.B. Unterrichtsplanung, Klassenführung, Umgang mit Unterrichtsstörungen und Störungsprävention, Umgang mit Heterogenität, Variablen der Unterrichtsqualität, Unterrichtsmethoden, -prinzipien und Sozialformen. Und wie oben gezeigt wurde, beeinflusst das epistemologische Überzeugungssystem eines Lehramtsstudierenden zu diesem erziehungswissenschaftlichen Wissen, wie tief diese Inhalte im Studium verstanden werden, was davon gelernt wird und was davon wie in die Unterrichtspraxis überführt wird (Buehl & Fives, 2016; Yadav & Koehler, 2007). Demnach ist es dringend nötig, ein quantitatives Erhebungsinstrument zu entwickeln, das es ermöglicht, die epistemologischen Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen empirisch zugänglich zu machen.

Bei der Erhebung ist jedoch dringend zu berücksichtigen, dass die epistemologischen Überzeugungen im Kontext aktiviert werden (Hofer, 2005, 2006b; Muis et al., 2006). Im Hinblick auf Lehramtsstudierende spielt der akademische und instruktionale Kontext eine besondere Rolle, da sich die Studierenden im akademischen Kontext befinden und das Lernen im Studium immer in einem speziellen instruktionalen Kontext stattfindet. Das bedeutet, diese beiden Kontexte beeinflussen sehr stark, welche epistemologischen Überzeugungen in den Lehramtsstudierenden aktiviert werden und demnach auch erhoben werden können. Um möglichst sicherzustellen, dass die befragten Lehramtsstudierenden die „richtigen“ epistemologischen Überzeugungen aktivieren, d.h. diejenigen, die hier von Interesse sind, wird der instruktionale Kontext in dieser Instrumententwicklung eingrenzt. Ganz

konkret wird sich die Erhebung auf die epistemologischen Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen und Wissensprozess aus Studientexten beziehen. Dieser Teilbereich wurde gewählt, da das selbstständige Lesen und Lernen aus Texten im Sinne des selbstgesteuerten Lernens ein wichtiger Aspekt des erziehungswissenschaftlichen Lehramtsstudiums ist und, wie oben dargelegt, die disziplinspezifischen epistemologischen Überzeugungen die Textverarbeitung und das Textverstehen stark beeinflussen (z.B. Bråten & Strømsø, 2004; Kardash & Noel, 2000; Schommer et al., 1992; Schommer & Walker, 1995; Schommer-Aikins, 2004; Schraw et al., 2002).

Aus dieser Synthese des theoretischen Hintergrundes leitet sich die übergeordnete Zielsetzung dieser Arbeit – die Testentwicklung – ab, aus der sich die zentrale Fragestellung ergibt. Die Zielsetzung lässt sich in spezifische Interpretationen der Testergebnisse, sogenannte Testwertinterpretationen, überführen. Daraus ergibt sich wiederum der konkrete Nutzen im Sinne des angestrebten Einsatzes des zu entwickelnden Tests und es lassen sich Grundannahmen ableiten, die für die Validierung der Testwertinterpretationen geprüft werden müssen.

# 3 Zentrale Zielsetzung, Fragestellung und Grundannahmen

Das Ziel dieser Arbeit ist die Entwicklung eines Instruments auf Basis der klassischen Testtheorie, das die mehrdimensionale und disziplinspezifische Erfassung epistemologischer Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen aus Studientexten erlaubt. Das resultierende Erhebungsinstrument wird **StEB (students' epistemological beliefs) Inventar** genannt.

Damit verbunden ist der angestrebte Einsatz (der Nutzen) des Tests. Das StEB Inventar soll der mehrdimensionalen Ermittlung der Ausprägungen der epistemologischen Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten dienen und sich folglich dazu eignen, die bei Lehramtsstudierenden aktuell vorherrschenden Ausprägungen in den epistemologischen Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten in den verschiedenen Dimensionen zu ermitteln. Der Einsatz des StEB Inventars soll darüber hinaus sowohl einen intraindividuellen als auch einen interindividuellen Vergleich der mehrdimensionalen Ausprägungen epistemologischer Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten ermöglichen. Das Inventar soll folglich dazu genutzt werden können, Aussagen darüber zu treffen, ob ein Lehramtsstudierender im Vergleich zu einem anderen über weniger sophistizierte (niedrige Merkmalsausprägung) oder aber sophistiziertere (höhere Merkmalsausprägung) epistemologische Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten (in den verschiedenen Dimensionen) verfügt. Über einen wiederholten Einsatz des StEB bei denselben Lehramtsstudierenden soll es außerdem möglich sein, individuelle Veränderungen in den Ausprägungen epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten (zu den verschiedenen Dimensionen) zu identifizieren. Diese verschiedenen intendierten Interpretationen der Testergebnisse und geplanten Arten der Testnutzung des StEB Inventars müssen in dieser Arbeit folglich validiert werden.

Zentral ist für diese Testentwicklung sowohl die theoretische als auch die methodische Fundierung. In Bezug auf die dargestellte Theorie zielt die Entwicklung einerseits darauf ab, die Mehrdimensionalität des Kontruktes (Conley et al., 2004; Hofer & Pintrich, 1997) abzubilden und darüber hinaus der Disziplinspezifität und Kontextspezifität (Muis et al., 2006)

epistemologischer Überzeugungen Rechnung zu tragen. Die Mehrdimensionalität wird umgesetzt, indem aus den theoretischen Kerndimensionen epistemologischer Überzeugungen die Dimensionen des zu entwickelnden Instruments abgeleitet werden. Der Wissensbereich, der im Fokus des Interesses steht, ist die Erziehungswissenschaft, genauer, die erziehungswissenschaftlichen Studienanteile des Lehramtsstudiums. Zwischen verschiedenen Schularten und Fächerkombinationen des Lehramtsstudiums sowie Fachsemestern wird nicht differenziert, da das resultierende Instrument möglichst bei Lehramtsstudierenden aller Fachgruppen und Schularten einsetzbar sein soll. Die Kontextspezifität wird durch den instruktionalen Kontext (vgl. Muis et al., 2006) der Studientexte zu erziehungswissenschaftlichen Studieninhalten noch expliziter herausgestellt, da Studientexte eine zentrale Quelle für den selbstgesteuerten Wissenserwerb in diesen Studienbereichen darstellen. Unter Studientext werden hier alle erziehungswissenschaftlichen Texte verstanden, die die Studierenden für Veranstaltungen und Prüfungen zum Lernen lesen. Das umfasst Texte, wie Forschungsartikel, Lehrbuchtexte und wissenschaftliche Beiträge, die von Hochschullehrenden in der Veranstaltung, im Seminar und auf Lernplattformen, wie ILIAS, Moodle und StudIP bereitgestellt werden.

Die zentrale Fragestellung dieser Arbeit, die sich aus der Zielsetzung ergibt, lautet:

**Fragestellung:** Lassen sich die epistemologischen Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen aus Studientexten in der Faktorstruktur des StEB, den theoretischen Annahmen entsprechend abbilden?

Das Instrument wird nach der klassische Testtheorie entwickelt, beschrieben und geprüft. Das heißt, sie liefert die Kriterien und Richtlinien, nach denen diese Instrumententwicklung umgesetzt und validiert wird. Folglich gliedert sich die Entwicklung übergeordnet in erstens die Entwicklung des Testentwurfs, zweitens die empirische Prüfung des Testentwurfs und drittens die Validierung (vgl. Bühner, 2011). Der Inhalt des ersten Testentwurfs, der auf einer starken theoretischen Basis ruht, wird zunächst über Expertenurteile gestützt (vgl. Döring & Bortz, 2016, S. 445 f.). Die Expertenbefragung zielt auf die Optimierung der Itemauswahl, des Iteminhalts und der -formulierung sowie der internen Struktur des Instruments, indem die Zuordnung der Items zu den Dimensionen geprüft wird. Nach der klassischen Testtheorie eignen sich zur empirischen Prüfung des Testentwurfs die exploratorische Faktorenanalyse (EFA) und die konfirmatorische Faktorenanalyse (CFA) (vgl. Brown, 2015, S. 20; Bühner, 2011). Nach idealtypischem Ablauf wird die EFA in einem ersten Schritt anhand einer ersten Stichprobe dazu eingesetzt, ungeeignete Items und Faktoren zu entfernen und die Faktorstruktur des Testentwurfes zu ermitteln (ebd.). Daraufhin wird diese gefundene Faktorstruktur in einem zweiten Schritt anhand einer neuen, größeren Stichprobe unter Einsatz des modellgenerierenden Ansatzes der CFA (vgl. Jöreskog, 1993) geprüft und validiert (vgl. Brown, 2015; Bühner, 2011; Fabrigar et al., 1999; R. B. Kline, 2016, S. 198). Die Objektivität, Reliabilität und Validität stellen die zentralen Gütekriterien der Entwicklung dieses Tests dar.

Diese Arbeit stützt sich auf das aktuelle Verständnis des Validitätsbegriffs nach den *Standards for Educational and Psychological Testing* der AERA, APA und NCME (2014), der im Rahmen der Beschreibung des methodischen Hintergrunds dieser Arbeit in Abschnitt 4.1 (ab S. 147) vorgestellt wird. Der Validierungsprozess ist zentral für diese Testentwicklung, da er zeigt, ob die intendierte Interpretation der Testergebnisse angemessen ist und demnach, ob der Test seinen beabsichtigten Nutzen erfüllen kann (AERA et al., 2014, vgl.).

Im Folgenden werden nach dem aktuellen argumentationsbasierten Validierungsansatz (AERA et al., 2014; vgl. auch Hartig et al., im Druck; siehe Abschnitt 4.1, ab S. 147) zunächst die konkreten Testwertinterpretation(en) spezifiziert und die zugehörigen empirisch prüfbaren Grundannahmen abgeleitet, die im späteren Verlauf dieser Arbeit zur Validierung des StEB geprüft werden.

**Testwertinterpretationen des StEB Inventars und Grundannahmen zur Validierung**

Für das StEB Inventar setzt die Validierung wie gefordert am theoretischen Hintergrund an, indem das Zielkonstrukt auf Basis einer umfassenden Literaturrecherche und Sichtung der vorhandenen theoretischen Konzeptualisierungen und Erhebungsinstrumente eingegrenzt und spezifiziert wurde (siehe Kapitel 2.10 Synthese, S. 126). Der Validierungsprozess zieht sich durch die gesamte Entwicklung des StEB Inventars. Die spezifischen intendierten Testwertinterpretationen und in diesem Zusammenhang der geplante Einsatz des StEB Inventars wurden oben im Zuge der Zielsetzung bereits beschrieben. An dieser Stelle seien sie noch einmal explizit aufgeführt:

**Testwertinterpretation 1 – erklärender Schluss:** Die mit dem StEB Inventar erfassten Testwerte sind Indikatoren für die Dimensionen epistemologischer Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten. Die Höhe der Ausprägung spiegelt sich in den Testwerten wider (vgl. Abschnitt 4.2, ab S. 150).

**Testwertinterpretation 2 – bewertender Schluss:** Die Testwerte des StEB Inventars können dazu eingesetzt werden, Personen mit geringerer Merkmalsausprägung, also weniger sophistizierten epistemologischen Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten, von Personen mit höherer Merkmalsausprägung, also sophistizierteren epistemologischen Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten, zu unterscheiden (vgl. Abschnitt 4.2, ab S. 150).

Zur Validierung dieser Testwertinterpretationen muss *erstens* gezeigt werden, dass die Testwerte sich als Maß für das theoretisch angenommene Konstrukt interpretieren lassen. Das Konstrukt epistemologischer Überzeugungen ist, wie im theoretischen Hintergrund dargestellt, *zweitens* mehrdimensional (Hofer, 2004a; Hofer & Pintrich, 1997; TIDE, Muis et al., 2006). Die theoretisch angenommene Binnenstruktur der epistemologischen Überzeugungen müsste sich folglich in der Faktorstruktur des StEB Inventars wiederfinden. Die Faktorstruktur sollte die Teilaspekte des übergeordneten Konstrukts reflektieren und im Einklang mit dem Theoriegefüge stehen. Aus diesem Theoriegefüge lässt sich *drittens* auch

auf die zwischen den verschiedenen Überzeugungsdimensionen vorherrschenden Beziehungen schließen, die sich im StEB Inventar ebenfalls zeigen müssten. Wenn es dem StEB gelingt, das Konstrukt einschließlich seiner Teilaspekte abzubilden, müssten sich zwischen den Überzeugungsdimensionen Zusammenhänge zeigen, da die einzelnen Überzeugungsdimensionen verschiedene Teilaspekte des übergeordneten Konstrukts epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten abbilden. Diese Zusammenhänge sollten jedoch moderat ($r < .40$) hoch sein, da das StEB Inventar für sich beansprucht, verschiedene Teilaspekte epistemologischer Überzeugungen zu erfassen und nicht mehrfach die gleichen.

Epistemologische Überzeugungen beeinflussen *viertens*, aufgrund ihrer Wirkung als Filter beim Wahrnehmen von Lernaufgaben, über die Auswahl von Lernstrategien den individuellen Lernprozess (Hofer, 2001; Hofer & Pintrich, 1997; Schommer-Aikins, 2004). Wenn das StEB Inventar also seinem angestrebten Nutzen entspricht und die Dimensionen epistemologischer Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten erfasst, müssten sich theoriekonforme Zusammenhänge zwischen den Skalen des StEB und den Lernstrategieskalen zeigen lassen, die in der Stärke ihrer Zusammenhänge vergleichbar mit den moderaten Zusammenhängen sind, welche sich in vergangenen Studien bereits vielfach empirisch zeigen ließen. Sophistiziertere Ausprägungen epistemologischer Überzeugungen (in den verschiedenen Dimensionen) müssten mit tiefenorientierten Lernstrategien (Elaborationsstrategien, metakognitive Lernstrategien und Organisationsstrategien) und weniger sophistizierte Ausprägungen epistemologischer Überzeugungen mit der oberflächenorientierten Wiederholungsstrategie einhergehen. Die Gültigkeit der angestrebten Nutzung kann demnach über die Prüfung der Verbindungen zwischen epistemologischen Überzeugungen und Lernstrategien geprüft werden.

Das Konstrukt epistemologischer Überzeugungen ist der Theorie zufolge kontextabhängig, da die individuellen epistemologischen Überzeugungen sich in Abhängigkeit vom akademischen und instruktionalen Kontext entwickeln (vgl. Kapitel 2.6, TIDE, Muis et al., 2006; Hofer, 2004a; Hofer & Pintrich, 1997). Wenn es dem StEB gelingt, diese Kontextabhängigkeit abzubilden, müsste sich demnach *fünftens* auch für die epistemologischen Überzeugungen zu erziehungswissenschaftlichem Wissen interindividuelle Unterschiede zwischen Lehramtsstudierenden der verschiedenen Hochschulstandorte finden lassen. Da nicht nur die spezifische Universität sondern auch der Studiengang einen Kontext zur Entwicklung epistemologischer Überzeugungen darstellt, ist *sechstens* davon auszugehen, dass interindividuelle Unterschiede zwischen Lehramtsstudierenden verschiedener Schularten vorliegen.

Für das Zielkonstrukt disziplinspezifischer epistemologischer Überzeugungen wird auf Basis der Theorie *siebtens* eine Weiterentwicklung angenommen (Bendixen & Rule, 2004; Hofer, 2004a; Hofer & Pintrich, 1997; TIDE, Muis et al., 2006), die auf den Einfluss von Bildungserfahrungen in der betreffenden Disziplin zurückzuführen ist. Nach den Erkenntnissen zur Entwicklung epistemologischer Überzeugungen (Chandler et al., 2002; King & Kitchener, 1994; Perry, 1970) entwickeln sich epistemologische Überzeugungen im Verlauf der Adoleszenz und des jungen Erwachsenenalters weiter (vgl. Kapitel 2.3, ab S. 51). Wie diese Entwicklung ausgelöst wird, beschreibt das IM Modell (Bendixen & Rule, 2004). Auch nach

dem TIDE Modell (Muis et al., 2006) entwickeln sich die akademischen epistemologischen Überzeugungen ab Eintritt in das Bildungssystem und mit zunehmenden akademischen Erfahrungen immer weiter (vgl. Kapitel 2.6, ab S. 75). Zusätzlich enwickeln sich über die akademischen Erfahrungen in den verschiedenen Fächern und Studiengängen disziplinspezifische epistemologische Überzeugungen, und auch von diesen wird angenommen, dass sie sich über die Zeit weiterentwickeln. Wenn es dem StEB gelingt, das Konstrukt abzubilden, müsste sich demnach auch für die epistemologischen Überzeugungen zu erziehungswissenschaftlichem Wissen über die Zeit eine Weiterentwicklung in Richtung sophistizierterer Überzeugungen zeigen lassen. Folglich müssten Unterschiede in den Ausprägungen epistemologischer Überzeugungen zwischen Lehramtsstudierenden hoher und niedriger Fachsemester zu finden sein. Auf Basis der Theorie wird für die epistemologischen Überzeugungen *achtens* angenommen, dass die primäre fachliche Ausrichtung eines Lernenden einen Einfluss auf die akademischen epistemologischen Überzeugungen und auch die anderen disziplinspezifischen epistemologischen Überzeugungen dieser Person hat und dieser Einfluss mit zunehmender Sozialisation in der spezifischen Disziplin weiter zunimmt (TIDE, Muis et al., 2006). Daraus müsste bei gültiger Interpretation der Testwerte des StEB Inventars folgen, dass bei Studienanfängern (in niedrigen Fachsemestern) mit unterschiedlicher fachlicher Ausrichtung (sozialwissenschaftlich, geisteswissenschaftlich und naturwissenschaftlich) keine Unterschiede in den epistemologischen Überzeugungen zum erziehungswissenschaftlichen Wissen zu finden sein dürften, bei fortgeschrittenen Lehramtsstudierenden (höherer Semester) jedoch zwischen den Lehramtsstudierenden unterschiedlicher fachlicher Ausrichtung (sozialwissenschaftlich, geisteswissenschaftlich und naturwissenschaftlich) Unterschiede in den epistemologischen Überzeugungen zum erziehungswissenschaftlichen Wissen auftreten müssten.

Die beiden Testwertinterpretationen (T1 und T2) werden folglich auf Basis der folgenden abgeleiteten Grundannahmen auf Gültigkeit geprüft:

**Grundannahme 1:** Die Antworten im StEB (Testwerte) lassen sich kausal auf verschiedene Überzeugungsdimensionen zurückführen.

**Grundannahme 2:** Das System resultierender Überzeugungsdimensionen zum erziehungswissenschaftlichen Wissen aus Studientexten ist mit der theoretischen Konzeptionalisierung nach Hofer und Pintrich (1997) vereinbar.

**Grundannahme 3:** Zwischen den Überzeugungsdimensionen des StEB liegen moderate Zusammenhänge vor, da die Überzeugungsdimensionen verschiedene Teilaspekte des übergeordneten Konstrukts *epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten* abbilden.

**Grundannahme 4:** Die epistemologischen Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten weisen moderate Zusammenhänge mit der berichteten präferierten Lernstrategieauswahl beim Lernen mit erziehungswissenschaftlichen Studientexten auf, wobei sophistiziertere Ausprägungen epistemologischer Überzeugungen (in den verschiedenen Dimensionen) mit tiefenorientierten Lernstrategien (Elaborationsstrategien, metakognitive Lernstrategien und Organisationsstrategien) einhergehen und weniger sophistizierte Ausprägungen epistemologischer Überzeugungen mit der oberflächenorientierten Wiederholungsstrategie einhergehen.

**Grundannahme 5:** Es bestehen Unterschiede in den Ausprägungen epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten zwischen Lehramtsstudierenden verschiedener Hochschulstandorte.

**Grundannahme 6:** Es bestehen Unterschiede in den Ausprägungen epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten zwischen Lehramtsstudierenden der Studiengänge Realschul-, Grundschul- und Sekundarschul- sowie Mittelschul- und Gymnasiallehramt (Schulart).

**Grundannahme 7:** Es bestehen Unterschiede in den Ausprägungen epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten zwischen Lehramtsstudierenden hoher und niedriger Fachsemester.

**Grundannahme 8:** Es bestehen Unterschiede in den Ausprägungen epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten zwischen Lehramtsstudierenden hoher Fachsemester mit sozialwissenschaftlicher, naturwissenschaftlicher und geisteswissenschaftfachlicher Ausrichtung. Für Lehramtsstudierende niedriger Fachsemester mit sozialwissenschaftlicher, naturwissenschaftlicher und geisteswissenschaftfachlicher Ausrichtung finden sich jedoch keine interindividuellen Unterschiede in den Ausprägungen epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen.

Zur Prüfung dieser acht Annahmen wird auf drei Evidenzquellen, nämlich den Testinhalt, die interne Teststruktur sowie die Beziehung der Testwerte zu anderen Variablen zurückgegriffen (für eine Beschreibung siehe Kap. 4.1 ab S. 147). Dabei werden Methoden eingesetzt, die im Zusammenhang mit dem alten Validierungsverständnis aus der Inhalts- und Konstruktvalidierung bekannt sind (für Inhaltsvalidität siehe S. 148 und Konstruktvalidität ab S. 148). Der Testinhalt wird im Rahmen einer Expertenbefragung (siehe Abschnitt 5.3, S.179) analysiert. Die Grundannahmen 1 und 2 beziehen sich auf die interne Teststruktur und lassen sich über die konfirmatorische Faktorenanalyse prüfen, deren Ergebnisse in Abschnitt 7.5.2 ab Seite 239 berichtet werden. Auch die Annahme 3 zu den Zusammenhängen zwischen den StEB-Dimensionen untereinander kann über die Ergebnisse der konfirmatorischen Faktorenanalyse getestet werden. Zusätzlich lassen sich korrelative Zusammenhänge ermitteln. Zur Prüfung der Annahme 4 zur Beziehungen zwischen den Testwerten des StEB

und den Testwerten der Lernstrategien wurde bereits im Rahmen der Synthese des theoretischen Hintergrundes ein nomologisches Netz aufgestellt (siehe Abb. 2.18, S. 130), das die angenommene Zusammenhangsstruktur abbildet. Die Grundannahme kann geprüft werden, indem die in der Validierungsstichprobe empirisch vorliegende Zusammenhangsstruktur mit der theoretisch angenommenen Zusammenhangsstruktur abgeglichen wird. Die Annahmen 5 bis 8 werden getestet, indem Mittelwertsunterschiede in StEB-Skalen zwischen Gruppen mittels MANOVAS geprüft werden. In Kapitel 8 ab Seite 273 werden die Prüfung all dieser Grundannahmen dargestellt, die Evidenzen gesammelt sowie zusammenfassend beurteilt.

# 4 Methodische Basis der Instrumentenentwicklung

Nachdem in den vorangegangenen Kapiteln der theoretische Hintergrund dieser Arbeit sowie die konkrete Zielsetzung vorgestellt und aus dieser die Fragestellung abgeleitet wurde, dient dieses Kapitel dazu, den methodischen Hintergrund zu skizzieren. Über die Darstellung der grundsätzlichen Logik der quantitativen Erfassung von Konstrukten im Rahmen von psychologischen Tests wird zur klassischen Testtheorie (KTT) hingeleitet. Deren Grundannahmen sowie die verschiedenen Messmodelle werden im Anschluss knapp aufgeführt, bevor die Testgütekriterien Objektivität, Reliabilität und Validität beschrieben werden. Im Zusammenhang mit der Validität wird zunächst auf das alte Validitätsverständnis eingegangen, nach dem verschiedene Validitätsarten unterschieden werden, bevor der aktuelle, argumentationsbasierte Validierungsansatz vorgestellt wird, nach dem angestrebte Testwertinterpretationen abgesichert werden. Dabei wird auf die Spezifikation von Testwertinterpretationen, das Aufstellen empirisch prüfbarer Grundannahmen, das Sammeln empirischer Evidenzen für und gegen die Grundannahmen sowie die zusammenfassende Beurteilung der Evidenzen eingegangen. Im Anschluss wird der idealtypische Ablauf einer Instrumentenentwicklung nach der KKT unter Einsatz der faktorenanalytischen Verfahren skizziert, indem zunächst die Teilschritte der exploratorischen Faktorenanalyse und anschließend die der konfirmatorischen Faktorenanalyse aufgezeigt werden.

Psychologische Tests dienen der quantitativen Erfassung von Fähigkeiten, Fertigkeiten, Eigenschaften und Zuständen von Personen (vgl. Bühner, 2011, S. 31 ff.). Meist handelt es sich in der Psychologie bei den zu erfassenden Merkmalen um Konstrukte, d.h. nicht direkt beobachtbare, sogenannte latente Variablen, wie beispielsweise Intelligenz oder epistemologische Überzeugungen. Über den Umweg von manifesten Variablen sollen latente Variablen messbar gemacht werden. Manifeste Variablen sind beobachtbares Verhalten, andere direkt messbare Informationen über eine Person (z.B. psysikalische oder physiologische) und auch Selbstauskünfte von Personen in Form von Itemantworten. Es wird angenommen, dass die zugrundeliegende latente Variable in ihrer spezifisch vorliegenden Merkmalsausprägung das manifeste Verhalten oder die Selbstauskünfte der Personen beeinflusst. Aus diesem Grund kann von der latenten Variable (z.B. epistemologische Überzeugung zu erziehungswissenschaftlichem Wissen und dem Wissensprozess) auf die manifesten Variablen, zum Beispiel ein bestimmtes Antwortverhalten in Items zur Selbstauskunft in einem Fragebogen geschlossen werden. Die manifesten Variablen dienen also als reflexive Indikatoren für die latente Variable (vgl. Bühner, 2011, S. 32 ff.). Das bedeutet, dass mit einer

Änderung in der Merkmalsausprägung immer auch eine Änderung der Ausprägung aller Variablen einhergehen muss. Die Messung soll demnach eindimensional sein (Eindimensionalitätsannahme), d.h. nur eine latente Variable soll ursächlich für die Ausprägung der Itembeantwortung sein. Damit der Schluss von der latenten auf die manifesten Variablen zulässig ist, muss die Eindimensionalitätsannahme bestand haben, die Konstrukte müssen empirisch gut erforscht sein und auf gesicherten Theorien basieren. Meist werden mehrere manifeste Variablen benötigt, um ein latentes Merkmal in seinen verschiedenen Facetten genau genug abzubilden. Soll die Messung über Selbstauskunft erfolgen, muss deshalb eine Sammlung von Items entwickelt werden, die das zu messende Merkmal abbilden (Bortz & Döring, 2006).

Als Basis für den kausalen Schluss von latenten auf manifeste Variablen dienen Testtheorien, wie beispielsweise die klassische Testtheorie (KTT) oder die probabilistische Testtheorie (PTT). Testtheorien liefern Erklärungen über den Zusammenhang zwischen Testverhalten und dem zu erfassenden Merkmal (Rost, 2004, S. 21). Sie definieren Kriterien der Qualität einer Messung (Gütekriterien) sowie aus welchen Bestandteilen die Messwerte bestehen (Bühner, 2011, S. 37). Wissenschaftliche Routineverfahren zur quantitativen Erfassung einer Merkmalsausprägung (einer Eigenschaft, Fähigkeit oder eines Zustands) oder aber qualitativen Zuordnung von Personen zu Personenklassen (Rost, 2004), die auf geprüften Theorien basieren, nach der KTT oder der PTT konstruiert wurden und den jeweiligen Gütekriterien genügen, werden *psychometrische Tests* genannt (Bühner, 2011, S. 38 f.).

Der Großteil der Testverfahren zu Erfassung epistemologischer Überzeugungen basiert bisher auf der klassischen Testtheorie. Sie stellt auch die Grundlage für die Instrumententwicklung in der vorliegenden Arbeit dar. Mittels der Kriterien der klassischen Testtheorie kann festgestellt werden, wie gut die einzelnen Items einer Skala das zu messende Merkmal repräsentieren. Sie definiert, aus welchen Bestandteilen sich die Messwerte zusammensetzen und liefert die Basis für die Definition der Messgenauigkeit und ihrer Anwendungen (Bühner, 2011, S. 37, 41). „Kennzeichnend für eine Testtheorie sind ihre Annahmen über die gemessenen Testwerte“ (Bortz & Döring, 2006, S. 193). Die folgenden Axiome stellen die Basis der klassischen Testtheorie dar, da sie die Grundannahmen über die gemessenen Testwerte ausdrücken (Bortz & Döring, 2006).

**Axiom 1** „Das Testergebnis (Score: X) setzt sich additiv aus dem wahren Wert (True Score: T) und dem Messfehler (Error Score: E) zusammen: $X = T + E$

**Axiom 2** Bei wiederholten Testanwendungen kommt es zu einem Fehlerausgleich, das heißt der Mittelwert $\mu_i$ des Messfehlers ist Null: $\mu(E) = 0$. Der Mittelwert mehrerer unabhängiger Messungen an demselben Untersuchungsobjekt ist folglich messfehlerfrei und repräsentiert den wahren Wert: $\mu(X) = \mu(T) + \mu(E) = T + 0 = T$. Da die wahre Merkmalsausprägung sich bei wiederholten Messungen an demselben Untersuchungsobjekt nicht ändert, gilt $\mu(T) = T$.

**Axiom 3** Die Höhe des Messfehlers ist unabhängig vom Ausprägungsgrad des getesteten Merkmals, d.h. wahrer Wert und Fehlerwert sind unkorreliert: $\rho_{T,E} = 0$

**Axiom 4** Die Höhe des Messfehlers ist unabhängig vom Ausprägungsgrad anderer Persönlichkeitsmerkmale $(T') : \rho_{T',E} = 0$

**Axiom 5** Die Messfehler verschiedener Testanwendungen (bei verschiedenen Personen oder Testwiederholungen bei einer Person) sind voneinander unabhängig, das heißt, die Fehlerwerte sind unkorreliert: $\rho_{E_1,E_2} = 0$.“ (siehe Bortz & Döring, 2006, S. 194)

Sogenannte Messmodelle der klassischen Testtheorie beschreiben die Beziehung zwischen den manifesten Variablen (Items) und der latenten Variable (Merkmal) (Bühner, 2011). Von dieser Beziehung hängt ab, welche Reliabilitätsschätzung angemessen ist. Die fünf bekanntesten Messmodelle sind nach Bühner (2011):

- das Modell paralleler Messung,
- das Modell essentiell paralleler Messung,
- das Modell tau-äquivalenter Messung,
- das Modell essentiell tau-äquivalenter Messung,
- das Modell tau-kongenerischer Messung.

Jedes dieser fünf Messmodelle gibt bestimmte Bedingungen vor, die die Itemmittelwerte (Itemschwierigkeit), die Itemfehlervarianzen (Messgenauigkeit der Items) und Ladungen der Items (Itemtrennschärfe) betreffen. Vor der Reliabilitätsanalyse muss also identifiziert werden, wie die Beziehung zwischen den Items und dem entsprechenden zu messenden Merkmal am besten zu beschreiben ist. Die Modelle unterscheiden sich in der Strenge ihrer Annahmen. Das Modell paralleler Messung gibt die stärksten Annahmen bezüglich der Itemmittelwerte (Itemschwierigkeit), der Itemfehlervarianzen (Messgenauigkeit der Items) und der Itemladungen (Itemtrennschärfe) vor. Nach der Reihe der Aufzählung werden die Modelle immer weniger restriktiv. Sind die Annahmen eines restriktiveren Modells erfüllt, können auch die weniger restriktiven Modelle, die weniger Bedingungen stellen, angenommen werden. In die andere Richtung gilt dies jedoch nicht (Bühner, 2011). Im Rahmen der Messmodellprüfung in Kapitel 7.6.1 (S. 263) wird im einzelnen auf die Bedingungen der verschiedenen Messmodelle eingegangen.

Allen Modellen gemeinsam ist die Annahme der statistischen Eindimensionalität. Damit der Schluss von der latenten auf die manifesten Variablen zulässig ist, muss die Messung eindimensional sein (vgl. Bühner, 2011, S. 31). Mit Eindimensionalität ist gemeint, dass die Ausprägung der manifesten Variablen, beispielsweise das Antwortverhalten einer Person, klar auf die Ausprägung eines einzigen zugrundeliegenden Konstrukts zurückzuführen sein muss. Wenn die Items eines Test mehrere latente Variablen messen, also nicht eindimensional sind, sind Testergebnisse nicht eindeutig interpretierbar, da die Ursache uneindeutig ist (ebd.) Dies lässt sich folgendermaßen darstellen: Da alle manifesten Variablen (Items) einer Skala dieselbe latente Variable messen, müssen sie miteinander korrelieren. Betrachtet man allerdings nur die Testergebnisse einer Personengruppe mit derselben Ausprägung in der latenten Variable, d.h. die latente Variable hat einen konstanten Wert, sollten

diese Personen bei einer exakten Messung alle das gleiche Antwortverhalten in den Items zeigen. Da keine Messung exakt ist, trifft dies natürlich nicht zu und es liegen Abweichungen im Antwortverhalten vor (Messfehler). Diese Messfehler sollten allerdings zufällig, d.h. unkorreliert sein. Sind sie es nicht, d.h. liegen Korrelationen zwischen den systematischen Messfehlern vor (= Residualkorrelationen), ist davon auszugehen, dass eine zweite latente Variable einen Einfluss auf das Antwortverhalten ausübt, denn da die ursprünglich interessierende latente Variable konstant gehalten wurde, d.h. keine Variation hat, kann sie die Unterschiede im Antwortverhalten nicht verursacht haben. Sind die Korrelationen zwischen den Items tatsächlich nur auf eine latente Variable zurückzuführen, spricht man von Items mit lokaler Unabhängigkeit (ebd.). Trotz der Eindimensionalitätsannahme kann ein Test natürlich verschiedene latente Variablen erheben, beispielsweise im Sinne von verschiedenen Subskalen eines übergeordneten Konstrukts. Für jede Dimension müssen in diesem Fall dann getrennte Testwerte berechnet und Reliabilitätsschätzungen vorgenommen werden (Bortz & Döring, 2006, S. 221).

## 4.1 Testgütekriterien

Über die drei zentralen Testgütekriterien Objektivität, Reliabilität und Validität ist es möglich, die Qualität eines entwickelten Tests einzuschätzen (Döring & Bortz, 2016). Gleichzeitig müssen diese Gütekriterien den gesamten Prozess der Instrumententwicklung von Anfang an informieren (AERA et al., 2014, vgl.).

Die Objektivität beschreibt das Ausmaß der Unabhängigkeit der Testanwendung vom Testanwender (vgl. Bortz & Döring, 2006, S. 195). Das bedeutet, dass unterschiedliche Testanwender, die den Test bei der gleichen Person durchführen, zu demselben Ergebnis kommen müssen. Man unterscheidet dementsprechend die Durchführungsobjektivität, Auswertungsobjektivität und Interpretationsobjektivität. Um das Kriterium der Durchführungsobjektivität zu erfüllen, muss das Testergebnis unabhängig vom Versuchsleiter sein. Standardisierte Instruktionen helfen Versuchsleitern dabei, die Durchführungsobjektivität zu sichern. Auswertungsobjektivität ist gewährleistet, wenn die Auswertung der einzelnen Testitems unabhängig von der Person ist, die den Test auswertet. Das heißt, die Vergabe der Punkte je Item darf nicht vom Auswertenden beeinflusst werden, sondern muss standardisiert erfolgen. Die Auswertungsobjektivität steht in Verbindung mit der Itemformulierung und der Antwortbewertung, da ein Item einfacher auszuwerten ist, wenn die Art, wie es vom Probanden beantwortet werden muss, vorgegeben ist. Geschlossene Items, das heißt Items mit vorgegebenen Antwortalternativen, die angekreuzt werden (z.B. „trifft zu“, „nein“, „selten“), sind demnach einfacher auszuwerten als offene oder halboffene Items. Das Kriterium der Interpretationsobjektivität ist erfüllt, wenn die Interpretation der Testergebnisse unabhängig von der individuellen Interpretation des Testanwenders ist, indem er beispielsweise normierte Vergleichswerte nutzt (vgl. Bortz & Döring, 2006, S. 195).

Die Reliabilität ist ein Maß für die Messgenauigkeit. Ein Instrument ist reliabel, wenn es das zu messende Merkmal genau misst. Da das Testergebnis sich nach der klassischen Testtheorie aus der echten Merkmalsausprägung des zu messenden Merkmals und dem

Messfehler zusammensetzt, ist die Reliabilität umso höher, je kleiner der Messfehler ist. Es handelt sich bei der Reliabilität also um ein Maß dafür, wie stark eine Messung durch einen Messfehler verzerrt ist (Döring & Bortz, 2016, S. 442). Da keine Messung ohne Fehler ist, ist eine vollständige Reliabilität nicht zu erreichen (vgl. Bortz & Döring, 2006, S. 196 f.). Werden Reliabilitätskoeffizienten auf Basis der klassischen Testtheorie ermittelt, die Annahmen zum Messfehler und wahren Wert einer Messung aufstellt, handelt es sich immer um pauschale Genauigkeitsschätzungen der Testwerte (Döring & Bortz, 2016, S. 443). Nach der klassischen Testtheorie ist die Reliabilität von Messungen als Varianzverhältnis definiert. „Es handelt sich um den Anteil der Varianz der wahren Werte an der Varianz der beobachteten Werte." (Bühner, 2011, S. 147). Da die Varianzen für die Population aus der Stichprobe geschätzt werden müssen, handelt es demnach um eine Schätzung der Reliabilität (ebd.). Die Reliabilität eines Tests lässt sich über die Retestreliabilität, die Paralleltestreliabilität, die Testhalbierungsreliabilität oder die interne Konsistenz schätzen (vgl. Bortz & Döring, 2006, S. 196 f.). Der Reliabilitätskoeffizient bewegt sich zwischen 0 und 1, wobei nur eine Messung, die völlig frei von Messfehlern ist, einen Wert von 1 erreicht. Üblicherweise wird eine Reliabilität von über .80 als ausreichend und ab .90 als hoch eingestuft (Bühner, 2011, S. 81). Döring und Bortz (2016, S. 443) warnen allerdings zurecht davor, diese Einstufung unreflektiert anzuwenden, da von der Art des gemessenen Merkmals und möglicher methodischer Alternativen abhängig sei, wie die Reliabilität zu interpretieren ist.

Die Validität ist das wichtigste, der Objektivität und Reliabilität übergeordnete aber von diesen wiederum abhängige, Gütekriterium (vgl. AERA et al., 2014, S. 11; Hartig et al., im Druck), denn verallgemeinert gesagt, befasst es sich damit, ob ein Test das misst, was er zu messen vorgibt (vgl. Hartig, Frey & Jude, 2012; Hartig et al., im Druck). In den internationalen Teststandards wird die Validität eines Tests definiert als „the degree to which evidence and theory support the interpretation of test scores for proposed uses of tests" (AERA et al., 2014, S. 11). Die Validität bezieht sich demnach auf das Ausmaß, zu dem empirische Evidenz und theoretische Argumente die Interpretation der Testwerte für die beabsichtigten Verwendungen des Tests stützen (vgl. AERA et al., 2014, S. 11). Die Validität wird nicht mehr, wie es nach dem alten Validitätsbegriff üblich war, als Eigenschaft eines Tests betrachtet, sondern dient der Absicherung von intendierten Testwertinterpretationen (vgl. AERA et al., 2014, S. 11; Hartig et al., im Druck). Nach den internationalen Teststandards ist es „incorrect to use the unqualified phrase *the validity of the test*" (AERA et al., 2014, S. 11), da es für einen Tests unmöglich ist, die Gültigkeit aller denkbaren Anwendungsmöglichkeiten und Testwertinterpretationen zu zeigen. Gleichzeitig werden auch nicht mehr verschiedene Validitätsarten, wie die Inhaltsvalidität und Konstruktvalidität mit den Unterkategorien faktorielle, divergente und konvergente Validität, unterschieden, sondern sie wird als einheitliches Qualitätskriterium betrachtet, das über einen argumentationsbasierten Ansatz Informationen aus verschiedenen Evidenzquellen integriert (vgl. Hartig et al., im Druck). Die verschiedenen Validitätsarten sind jedoch noch insofern von Bedeutung, als dass die Methoden, die eingesetzt wurden, um Evidenz für die einzelnen Validitätsarten zu sammeln, nach wie vor eingesetzt werden (ebd.). Nach dem aktuellen Verständnis werden die Ergebnisse jedoch integriert dargestellt (vgl. Hartig et al., im Druck).

Aufgrund der andauernden Bedeutung der eingesetzten Methoden wird im Folgenden auf die einzelnen Validitätsarten und zugehörigen Prüfungsmethoden eingegangen, die nach dem früheren Validitätskonzept unterschieden wurden, und welche beispielsweise bei Döring und Bortz (2016) nach wie vor unterschieden werden.

**Inhaltsvalidität** Die Inhaltsvalidität als Teilaspekt der Validität befasste sich mit der Frage, ob die einzelnen Items des Tests das zu erfassende Konstrukt in seiner inhaltlichen Breite vollständig sowie inhaltlich richtig abbilden (Döring & Bortz, 2016, S. 446). Da Konstrukte komplex sind und aus verschiedenen Teilaspekten bestehen, müssen diese durch die Items inhaltlich auch abgedeckt werden. Die Inhaltsvalidierung setzte an der Theorie an. Denn um die verschiedenen Bedeutungsaspekte eines Zielkonstrukts inhaltlich abzudecken, müssen diese zunächst identifiziert werden. Im Rahmen einer Konzeptspezifikation muss das Konstrukt klar definiert werden. Es muss festgestellt werden, welche inhaltlichen Bedeutungsaspekte es umfasst und welche dagegen anderen anschließenden Konstrukten zuzuordnen sind. Das heißt, eine klare Abgrenzung ist erforderlich. Ist dies geschehen, können die einzelnen Testitems diesen verschiedenen inhaltlichen Aspekte des Konstrukts zugeordnet werden bzw. zu jedem Teilaspekt werden passende Items formuliert. Danach muss geprüft werden, ob alle relevanten Aspekte abgedeckt sind und ob die Items die theoretisch vorgegebene Bedeutung richtig widerspiegeln (Döring & Bortz, 2016, S. 446). Für die Inhaltsvalidierung steht kein Koeffizient zur Verfügung, sie muss vielmehr anhand der Theorie argumentativ vorgenommen werden und kann zusätzlich über Expertenurteile gestützt werden. Verbesserungsvorschläge der Fachexperten führen dann gegebenenfalls zu Anpassungen der Testentwicklung (Döring & Bortz, 2016, S. 445 f.).

**Konstruktvalidität** Unter Konstruktvalidität verstand man die Gültigkeit eines Tests im Bezug auf das zu messende Konstrukt. Das heißt, bei der Prüfung der Konstruktvalidität wird der Frage nachgegangen, ob der Test tatsächlich das Konstrukt (die Eigenschaft oder Fähigkeit) misst, das er zu messen vorgibt (Bühner, 2011, S. 63 f.). Ein Test ist konstruktvalide, wenn er inhaltlich und theoretisch begründet hypothesenkonform mit anderen theoretischen Konstrukten korreliert (Döring & Bortz, 2016, S. 446 f.). Dazu muss das Zielkonstrukt auf theoretischer und empirischer Basis genau definiert, abgegrenzt sowie die Art der Verbindung zu angrenzenden und überlappenden Konstrukten identifiziert werden. Wenn ein Netz aus gut fundierten Hypothesen darüber vorliegt, welche Zusammenhangsstrukturen (Richtung und Enge) das Zielkonstrukt mit anderen Konstrukten aufweist, idealerweise ein *nomologisches Netz* (Cronbach & Mehl, 1955), kann eine Konstruktvalidierung erfolgen, indem diese Zusammenhänge bivariat, multivariat (z.B. multiple Regressionsanalysen) oder mittels Strukturgleichungsmodellen getestet werden (Hartig et al., 2012, S. 153 f.).

Die Konstruktvalidität umfasste wiederum die konvergente, diskriminante und faktorielle Validität (Döring & Bortz, 2016, S. 445 f.).

**Konvergente Validität** Sie lag vor, wenn der zu „validierende Test mit eng verwandten Konstrukten bzw. konstruktnahen Verfahren relativ hoch korreliert (auf jeden Fall

höher als mit konstruktfremden Verfahren)“ (Döring & Bortz, 2016, S. 446). Aus der Theorie wurde dazu eine Zusammenhangsstruktur zwischen dem Zielkonstrukt und einem verwandten Konstrukt oder bestimmten manifesten Variablen abgeleitet, ein sogenanntes nomologisches Netz (Hartig et al., 2012). Die theoretisch angenommenen Zusammenhänge stellen Vorhersagen dar, die mit den empirischen Beobachtungen verglichen werden können (ebd.). Die Prüfung der empirischen Zusammenhangsstruktur kann mittels eines korrelativen Ansatzes und auch über Strukturgleichungsmodelle vorgenommen werden (Hartig et al., 2012, S. 158). Strukturgleichungsmodelle erlauben es „die Zusammenhangsstruktur mehrerer Variablen bei gleichzeitiger Kontrolle der Messfehler“ zu prüfen (Hartig et al., 2012, S. 158). Beim korrelativen Ansatz werden dagegen empirisch ermittelte, bivariate Korrelationen mit den theoretisch angenommenen Zusammenhängen verglichen (ebd.). Ein typischer Ansatz der Konstruktvalidierung besteht in der Untersuchung der korrelativen Beziehungen „zwischen den Testwerten X, deren Interpretation validiert werden soll, und einer anderen manifesten Variable Y. Bei Y kann es sich ebenfalls um einen Testwert, ein Verhaltensmaß oder um jedwede andere Personenvariable (z.B. Alter oder Geschlecht) handeln.“ (Hartig et al., 2012, S. 158)

**Diskriminante/divergente Validität** Sie war gegeben, wenn der zu validierende Test keine oder nur geringe Zusammenhänge zu benachbarten Konstrukten aufweist (Döring & Bortz, 2016, S. 446). Es musste dabei darauf geachtet werden, dass für die diskriminante Validierung nicht weit entfernte Konstrukte gewählt werden, zu denen von vornherein keine Verbindung zum Zielkonstrukt erwartet werden, sondern es sich tatsächlich um benachbarte Konstrukte aus demselben Gültigkeitsbereich (Leistungs- oder Persönlichkeitsbereich) handelte (Bühner, 2011, S. 80).

**Faktorielle Validität** Von faktorieller Validität als Teilaspekt der Konstruktvalidität war in Bezug auf ein Erhebungsinstrument auszugehen, wenn die Items, von denen angenommen wird, dass sie inhaltlich einen bestimmten Teilaspekt des Konstrukts abdecken, auch empirisch einen Faktor bilden. Bei mehrdimensionalen Tests sollten sich die Items also mittels Faktorenanalyse ihren inhaltlich passenden Subskalen zuordnen lassen. Bei einem eindimensionalen Test muss gezeigt werden, dass alle Items gemeinsam nur einen einzigen Faktor bilden. Dies kann mittels der exploratorischen oder konfirmatorischen Faktorenanalyse geprüft werden (vgl. Döring & Bortz, 2016, S. 446).

**Kriteriumsvalidität** Ein Test wurde als kriteriumsvalide betrachtet, wenn der Testwert „positiv mit inhaltlich korrespondierenden manifesten Merkmalen außerhalb der Testsituation“ (Döring & Bortz, 2016, S. 447) korrelierte. Je nach Zeitpunkt der Erhebung des Außenkriteriums unterschied man zwischen retrospektiver, konkurrenter und prognostischer Validität. Bei der retrospektiven Validierung wurde das Außenkriterium zeitlich vor dem zu validierenden Testwert erhoben, während es bei der konkurrenten (auch Übereinstimmungsvalidität) zum gleichen Messzeitpunkt und bei der prognostischen Validität (auch

Vorhersagevalidität) zeitlich danach erfasst wurde (Döring & Bortz, 2016, S. 447). Als Richtlinie zur Beurteilung der Kriteriumsvalidität eines Tests nennt Bühner (2011, S. 80 f.) eine Korrelation von $r > .20$. Die inkrementelle Validität gehörte ebenfalls zur Kriteriumsvalidierung eines Tests. Sie sagt aus, ob dieser neue Test enger mit einem Außenkriterium korreliert oder dieses besser vorhersagt, als bisher verfügbare Tests (Döring & Bortz, 2016, S. 447). Da eine Kriteriumsvalidierung nur umzusetzen ist, wenn ein valides und reliables Außenkriterium zur Verfügung steht (Döring & Bortz, 2016, S. 447), was bei der Erfassung von psychologischen Konstrukten selten der Fall ist, wird diese Art der Validierung weit weniger durchgeführt als beispielsweise die Konstruktvalidierung (Hartig et al., 2012).

## 4.2 Argumentationsbasierter Validierungsansatz

Nach dem aktuellen Verständnis von Validität, das einem argumentationsbasierten Ansatz folgt und die oben beschriebenen Validitätsarten nicht mehr unterscheidet, gliedert sich die Validierung in die folgenden vier Teilschritte, die insgesamt das sogenannte Validitätsargument – eine stringente Argumentation zur Stützung der Validität von bestimmten Testwertinterpretationen – bilden (vgl. Hartig et al., im Druck):

1. Testwertinterpretation(en) spezifizieren
2. Empirisch prüfbare Grundannahmen aufstellen
3. Empirische Evidenzen für und gegen die Grundannahmen sammeln
4. Zusammenfassende Beurteilung der Evidenzen für die Testwertinterpretation(en)

Im ersten Schritt wird spezifiziert und ausformuliert, wie das Ergebnis des Tests, d.h. der Testwert oder die Testwerte, interpretiert werden soll. Dazu muss zunächst das Zielkonstrukt eingegrenzt und definiert werden (vgl. Hartig et al., im Druck). Die Testwertinterpretation bezieht sich darauf, was aus den Testwerten gefolgert werden kann und betrifft die intendierte Verwendung des Tests. Nach Kane (2001) kann zwischen fünf typischen Testwertinterpretationen unterschieden werden: Bewertung, Verallgemeinerung, Extrapolation, Erklärung und Entscheidungsfindung (siehe Hartig et al., im Druck). Bei der Testwertinterpretation handelt es sich um einen *bewertenden Schluss*, wenn aufgrund der Testwerte von Personen Vergleiche bezüglich der jeweiligen Merkmalsausprägungen vorgenommen werden (ebd.). Ein Beispiel hierfür ist, wenn auf der Basis der Testwerte des StEB Inventars Aussagen darüber getroffen werden, ob Personen im Vergleich zu anderen über weniger sophistizierte oder aber sophistiziertere epistemologische Überzeugung verfügen. Vom *verallgemeinernden Schluss* spricht man, wenn eine getestete Person mit einem spezifischen Testwert (z.B. hoher Testwert) auch ähnliche Items im Sinne der vorliegenden Merkmalsausprägung (hoher Testwert, hohe Merkmalsausprägung) beantworten wird (z.B. bei Leistungstestaufgaben werden andere, ähnliche Aufgaben auch gelöst). Ein *extrapolierender Schluss* erfolgt, wenn auf Basis des spezifischen Testwerts Schlussfolgerungen für andere Bereiche

gezogen werden, z.B. wenn auf Basis der Testergebnisse eines Studieneignungstests auf den späteren Studienerfolg geschlossen wird. Bei der Testwertinterpretation handelt es sich um einen *erklärenden Schluss*, wenn „das Testergebnis mit Bezug auf Theorien [...] als Indikator für ein theoretisches Konstrukt" (Hartig et al., im Druck, S. 7) interpretiert wird, d.h. wenn aufgrund des Testwerts auf ein zugrundeliegendes Konstrukt geschlossen wird. Dies ist beispielsweise der Fall, wenn auf Basis der Testergebnisse des StEB Inventars mit Bezug zu dem theoretischen Modell von Hofer und Pintrich (1997) auf individuelle Merkmalsausprägungen in den Kerndimensionen epistemologischer Überzeugungen geschlossen wird. Die fünfte Art der Testwertinterpretation liegt vor, wenn aus dem Testergebnis eine *schlussfolgernde Entscheidung* abgeleitet wird (Hartig et al., im Druck, S. 7 f.) (z.B. wenn bei einem Einstellungstest das Ergebnis über die Vergabe der Stelle entscheidet).

Nachdem die Testwertinterpretationen präzise festgelegt wurden, werden daraus empirisch überprüfbare Grundannahmen aufgestellt, die für die jeweilige Testwertinterpretation bedeutsam sind und sie vollständig abdecken. Wichtig ist, dass jede einzelne der angestrebten Testwertinterpretation eine separate Validierung erfordert, also ihrer eigenen, prüfbaren Grundannahmen bedarf. Die Grundannahmen müssen potenziell falsizifierbar sein und es dürfen keine relevanten Grundannahmen fehlen. Mittels der Grundannahmen sollte prüfbar werden, ob der Test das zugrundeliegende Konstrukt hinreichend gut repräsentiert, insofern er es weder unterrepräsentiert (d.h. wesentliche Aspekte nicht abdeckt) oder aber konstruktirrelevante Aspekte misst (Hartig et al., im Druck).

Im Anschluss wird empirische Evidenz gesammelt, anhand derer die aufgestellten Grundannahmen entweder widerlegt oder aber gestützt werden können (ebd.). Als mögliche Evidenzquellen dienen nach AERA et al. (2014, S. 13 f.):

1. der Testinhalt,
2. der Antwortprozess,
3. Beziehungen der Testwerte zu anderen Variablen sowie
4. die interne Struktur des Tests.

Wird der *Testinhalt* bzw. die Beziehung zwischen Testinhalt und dem Zielkonstrukt als Evidenz herangezogen, dient eine genaue Spezifikation des Zielkonstrukts als Ausgangspunkt der Analysen. Logische/argumentative und empirische Analysen von z.B. Expertenurteilen über Itemformulierung, -inhalt und Repräsentativität der Items werden eingesetzt, um zu untersuchen, ob der Testinhalt den Inhaltsbereich des Zielkonstrukts adäquat repräsentiert (AERA et al., 2014, S. 14 f.). Hier zeigt sich deutlich die Nähe zum Begriff der Inhaltsvalidität (siehe oben S. 148). Der *Antwortprozess* kann als wichtige Evidenzquelle dienen, wenn die

Testinterpretation beispielsweise explizite Annahmen zu den spezifischen, bei der Testbeantwortung ablaufenden kognitiven Prozessen aufstellt. Sind solche Annahmen zum Antwortprozess relevant, müssen entsprechend empirische oder theoretische Analysen durchgeführt werden, um diese Annahmen zu prüfen (AERA et al., 2014, S. 14 f.). Häufig impliziert die angestrebte Nutzung des Tests bestimmte *Beziehungen der Zielkonstrukts zu anderen Variablen*. Zur Prüfung dieser Testwertinterpretation müssen demnach die Beziehungen zwischen dem Testwert und den entsprechenden Variablen, z.B. den Testwerten anderer Messinstrumente, untersucht werden (vgl. AERA et al., 2014, S. 16 f.). Die dem Zielkonstrukt zugrundeliegende Theorie kann beispielsweise bestimmte Zusammenhänge zwischen den Testwerten und einer anderen Variable implizieren, wenn eine bestimmte Testwertinterpretation gültig ist (Bsp.: Sophistizierte epistemologische Überzeugungen korrelieren positiv mit den Elaborationsstrategien.). Die zugrundeliegende Theorie kann beispielweise auch Gruppenunterschiede in einer bestimmten Variable zwischen Personen mit hohen und niedrigen Testwerten implizieren, wenn eine bestimmte Testwertinterpretation gültig ist (Bsp.: Personen mit sophistizierten Überzeugungen zur Sicherheit des Wissens setzen mehr tiefenorientierte Lernstrategien ein als Personen mit wenig sophistizierten Überzeugungen zur Sicherheit des Wissens.). Nach dem alten Validitätsverständnis beziehen sich diese Aspekte auf die konvergente und divergente Validierung (siehe oben Konstruktvalidität ab S. 148). Die *interne Struktur* des Tests dient als wichtige Evidenzquelle zur Prüfung von Annahmen darüber, ob die Beziehungen zwischen Testitems und Testdimensionen mit dem Konstrukt, auf dem die angenommene Testwertinterpretation basiert, korrespondieren (AERA et al., 2014, S. 16). Beispielsweise könnten aus der Theorie abgeleitete Annahmen dazu vorliegen, wie viele Teilaspekte das Zielkonstrukt des Tests aufweist und welche Zusammenhänge zwischen diesen Teilaspekten bestehen. Solche Annahmen ließen sich in Hypothesen zur Dimensionalität des Tests und zu vorliegenden Korrelationen zwischen diesen Dimensionen ausdrücken, die faktorenanalytisch prüfbar wären (vgl. AERA et al., 2014, S. 16). Die Prüfung der faktoriellen Validität als Teil der Konstruktvalidität aus dem alten Validitätsverständnis weist hohe Übereinstimmung zu diesem Aspekt auf (siehe oben ab S. 148). Im Validierungsprozess wird auf eine oder auch mehrere dieser vier Evidenzquellen Bezug genommen, um die Gültigkeit einer jeden angenommenen Testwertinterpretation zu rechtfertigen (Hartig et al., im Druck).

Den Abschluss der Validierung bildet eine zusammenfassende Beurteilung der Testwertinterpretation unter Berücksichtigung aller Evidenzen, die für oder gegen die verschiedenen Grundannahmen sprechen. Werden alle Grundannahmen bestätigt, stützt dies die Annahme, dass die angenommene Interpretation der Testergebnisse und der angestrebte Einsatz des Test vorläufig valide (gültig) sind. Wird eine Grundannahme jedoch widerlegt, stellt dies mindestens eine Einschränkung für die betreffende Testwertinterpretation und den intendierten Testeinsatz dar, wenn sie nicht sogar verworfen werden müssen (Hartig et al., im Druck).

## 4.3 Prüfung des Testentwurfs und Validierung mittels EFA und CFA

Mittels faktorenanalytischer Verfahren, der exploratorischen und konfirmatorischen Faktorenanalyse (EFA und CFA), wird die in den Daten liegende Faktorenstruktur identifiziert, die die Beziehungen zwischen den beobachteten Daten erklären kann. Es werden latente Variablen identifiziert, die die Zusammenhänge zwischen den manifesten Variablen erklären können (Fabrigar et al., 1999). Für die Testentwicklung nach der klassischen Testtheorie eignen sich die exploratorische Faktorenanalyse (EFA) und die konfirmatorische Faktorenanalyse (CFA) demnach zur empirischen Prüfung des Testentwurfs (vgl. Brown, 2015, S. 20; Bühner, 2011). Die EFA kann in einem ersten Schritt anhand einer ersten Stichprobe dazu eingesetzt werden, um ungeeignete Items zu entfernen und die Faktorstruktur des Testentwurfes zu ermitteln. Daraufhin kann diese gefundene Faktorstruktur in einem zweiten Schritt anhand einer neuen, größeren Stichprobe unter Einsatz der CFA geprüft werden (vgl. Brown, 2015; Bühner, 2011; Fabrigar et al., 1999; R. B. Kline, 2016, S. 198).

Im Folgenden wird in die exploratorische und konfirmatorischen Faktorenanalyse eingeführt sowie die Teilschritte beider Verfahren überblicksartig dargestellt. Für eine ausführlichere und genauere Darstellung der EFA sei auf Bühner (2011), Brown (2015, S. 10-34) und Fabrigar und Wegener (2012), für die CFA dagegen auf Brown (2015), R. B. Kline (2016) und ebenfalls Bühner (2011) verwiesen. R. B. Kline (2013) stellt die Unterschiede zwischen EFA und CFA dar.

### 4.3.1 Exploratorische Faktorenanalyse

Das Ziel einer exploratorischen Faktorenanalyse ist, eine Vielzahl von korrelierenden, manifesten Variablen (Items) auf einen kleinen Satz latenter Variablen (Faktoren) zu reduzieren, die einen möglichst großen Teil der Varianz der Items aufklären (Baglin, 2014; Bühner, 2011). Wie alle faktorenanalytischen Methoden zerlegt auch die EFA die standardisierte Varianz eines Indikators (Items) in die *common variance* (gemeinsame Varianz) und die Fehlervarianz (Uniqueness) (R. B. Kline, 2016, S. 190). Die *common variance* bezeichnet den Varianzanteil, Kommunalität $h^2$ genannt, den sich das Item mit den übrigen Items teilt, da es mit diesen zusammenhängt. Es handelt sich also um den gemeinsamen Varianzanteil, der durch die latenten Variablen (Merkmale) erklärt werden kann. Ein zugrundeliegendes Merkmal (z.B. Überzeugung zu Wissen) wird demnach als Ursache dafür gesehen, dass die Items untereinander Varianz teilen (R. B. Kline, 2016, S. 190). Die restliche Varianz, die nicht auf die latenten Variablen zurückzuführen ist, wird als *Uniqueness* bzw. Fehlervarianz ($1 - h^2$) bezeichnet. Sie besteht eigentlich aus der itemspezifischen Varianz, d.h. der Varianz eines Items, die auf die spezielle Itemcharakteristik zurückgeht und nicht durch Faktoren erklärbar ist sowie der Messfehlervarianz. Diese beiden werden im Rahmen der Faktorenanalyse jedoch nicht unterschieden (ebd.). Aus diesen verschiedenen Varianzanteilen, von denen nur die *common*

*variance* auf die Faktoren zurückzuführen ist, folgt, dass die Unterschiede in den Merkmalsausprägungen (Unterschiede in Überzeugung zu Wissen) nicht die gesamten Unterschiede in der Itembeantwortung erklären können (Bühner, 2011, S. 310).

„Das Ergebnis der Faktorenanalyse sind wechselseitig voneinander unabhängige Faktoren, die die Zusammenhänge zwischen den Variablen erklären." (Bortz & Schuster, 2010, S. 386). Die EFA stellt demnach eine primäre Methode der Testentwicklung dar. Sie eignet sich dazu, die Faktorstruktur eines Tests zu untersuchen und die Itemauswahl zu optimieren (Brown, 2015, S. 10). Anders als bei der CFA wird die Faktorstruktur hier ohne die Vorgabe einer faktorenanalytischen Hypothese analysiert, was deswegen von Vorteil ist, da sich mit der EFA die theoretische Struktur auch statistisch ohne modellierende Vorannahmen zeigen muss. Die EFA mit einem theoretischen A-Priori-Modell zu vergleichen, ist daher eingangs die konservativste Vorgehensweise: Das Modell entsteht nicht aus der Faktorstruktur der EFA, das Modell muss frei berechnet dem angenommenen theoretischen Modell entsprechen.

Startpunkt der EFA ist die Interkorrelationsmatrix der manifesten Variablen. Sie produziert je Faktor für jedes Item eine Faktorladung $\lambda$. Der Wert der Faktorladung reicht von -1 bis 1. Eine hohe Ladung eines Items auf einem Faktor (nahe 1 oder -1) drückt aus, dass die latente Variable mit dem Item stark zusammenhängt und dieses bei der inhaltlichen Interpretation des Faktors eine besondere Bedeutung hat. Eine niedrige Ladung auf einem Faktor drückt dagegen aus, dass die latente Variable mit dem Item schwach oder gar nicht zusammenhängt und es bei der inhaltlichen Interpretation des Faktors eine untergeordnete oder keine Bedeutung spielt. Die inhaltliche Interpretation der Faktoren erfolgt auf Basis der Inhalte der zugehörigen Items. Eine Einfachstruktur, d.h. eine hohe Ladung der Items auf jeweils nur einem Faktor (ohne Nebenladungen), erleichtert dabei die Interpretation der Faktoren (Bühner, 2011). Neben den Faktoren und Faktorladungen führt eine EFA zu den folgenden Ergebnissen: den Eigenwerten der Faktoren, den Kommunalitäten ($h^2$), der Ladungsmatrix und den Faktorwerten (Bühner, 2011). Der Eigenwert eines Faktors gibt an, wie viel Varianz dieser Faktor erklärt. Er ist die Summe der quadrierten Faktorladungen eines Faktors über alle Variablen. Die Kommunalität gibt dagegen an, wie viel Varianz die Variable (Item) an allen Faktoren erklärt. Die Kommunalität ist die Summe der quadrierten Ladungen einer Variablen über alle Faktoren. Der Faktorwert gibt die Ausprägung des entsprechenden Faktors (Merkmals) bei einem Individuum wieder und errechnet sich als Linearkombination der Faktorladungen mit den individuellen Ausprägungen der manifesten Variablen bei einer Person. Die Faktorwerte sind standardnormalverteilt (Bühner, 2011).

Die Durchführung einer EFA gliedert sich in die folgenden Arbeitsschritte (vgl. Bühner, 2011; Brown, 2015, S. 34):

- Prüfung der Voraussetzungen
- Wahl der geeigneten Extraktionsmethode
- Bestimmung der geeigneten Faktorenanzahl

- Faktorrotation
- Entfernen ungeeigneter Items und Interpretation der Faktoren
- Nach Itemreduktion erfolgt erneute Prüfung mittels EFA

Zur Prüfung der Voraussetzungen der EFA gehört es zunächst, sich darüber klar zu werden, ob eine EFA tatsächlich das richtige Verfahren darstellt oder sich vielmehr eine Hauptkomponentenanalyse (*Principal Component Analysis*, PCA)[1] eignet (vgl. Baglin, 2014; Fabrigar et al., 1999; Schmitt, 2011). Da sowohl die EFA als auch PCA dazu dienen, die Anzahl der in den Daten liegenden Dimensionen zu reduzieren, ist es wichtig, sich den Unterschied zwischen beiden bewusst zu machen. Mittels der Faktorenanalyse (EFA und CFA) wird die in den Daten liegende Faktorenstruktur identifiziert, die die Beziehungen zwischen den beobachteten Daten erklären kann, d.h., es werden latente Variablen identifiziert, die die Korrelationen zwischen den manifesten Variablen erklären können (Fabrigar et al., 1999). Die Faktorenanalyse eignet sich daher dazu, die Dimensionalität einer Skala zu erkunden und eine den beobachteten Daten zugrundeliegende latente Variable vorzuschlagen (Baglin, 2014, S. 2). Die PCA zielt dagegen darauf ab, eine große Anzahl miteinander zusammenhängender Variablen bei möglichst geringem Informationsverlust auf eine geringere Anzahl von Komponenten zu reduzieren. Gleichzeitig zielt die PCA nicht darauf ab, die den Daten zugrundeliegende Faktorstruktur der Population zu ermitteln (Baglin, 2014; Fabrigar et al., 1999). Im Falle einer Instrumententwicklung wie der vorliegenden, ist also klar eine Faktorenanalyse (EFA oder CFA) und nicht die PCA einzusetzen.

Der erste Einsatz des StEB soll zunächst mittels einer EFA ausgewertet werden, da es sich eben einerseits um den ersten Einsatz handelt und es sich bei den Skalen epistemologischer Überzeugungen bisher eher um tentative Modelle handelte (vgl. Pintrich, 2002). So kann zunächst die Faktorstruktur des Testentwurfes ermittelt und die Itemauswahl optimiert werden.

Generell gilt in Bezug auf die Faktorenanalyse, dass größere Stichproben zu stabileren Faktorlösungen führen (Bühner, 2011, S. 345). Bezüglich der erforderlichen Stichprobengröße für die Faktorenanalyse lassen sich verschiedene Richtlinien finden, die ein Verhältnis zwischen Item- und Versuchspersonenanzahl postulieren (für eine Übersicht siehe z.B. Bryant & Yarnold, 2001, vgl. auch Fabrigar et al., 1999). Überzeugender ist das von Mac Callum, Widaman, Zhang und Hong (1999) in Monte-Carlo-Studien abgeleitete Verhältnis zwischen Stichprobengröße und Kommunalitäten. Hier zeigte sich, dass bereits eine Stichprobengröße von $N = 60$ ausreichend sein kann, sofern die Kommunalitäten aller Items hoch sind ($h^2 > .60$) (vgl. Bühner, 2011; Fabrigar et al., 1999). Bei mittleren Kommunalitäten ($h^2 \approx .50$) wird bereits ein Stichprobenumfang von 100 bis 200 benötigt (vgl. Bühner, 2011; Fabrigar et al., 1999). Anhaltspunkte für die Abhängigkeiten zwischen Stichprobengröße, Itemkommunalitäten und Itemanzahl je Faktor geben auch Mundform, Shaw und Ke (2005); allerdings in einer Studie, die sich auf die Maximum-Likelihood-Faktorenanalyse

[1] Die PCA wird oft fälschlicherweise als Extraktionsmethode der Faktorenanalyse angesehen, gehört aber streng genommen nicht zur Familie der Faktorenanalysen (vgl. Schmitt, 2011 und Bühner, 2011, S. 309).

mit Varimax-Rotation bezogen hat. Trotzdem ist der Grundsatz übertragbar, dass die Stichprobe umso kleiner sein kann (bis minimal $N = 55$), je höher die Itemanzahl je Faktor und die Itemkommunalitäten sind (Bühner, 2011, S. 345). Jedem Faktor sollten mindestens drei bis fünf Items zugeordnet sein (vgl. Mac Callum et al., 1999; Velicer & Fava, 1998, und Bühner, 2011, S. 344). Die Abhängigkeiten zwischen Stichprobengröße, Itemkommunalitäten und Itemanzahl je Faktor nach Bühner (2011, S. 345) werden in Tabelle 4.1 dargestellt.

TABELLE 4.1: Stichprobengröße für EFA in Abhängigkeit von Kommunalität für ML-Methode nach Bühner (2011, S. 345), in eigener Darstellung

| | $h^2$ (.60 - .80) | | $h^2$ (.40 - .60) | | $h^2$ (.20 - .40) | |
|---|---|---|---|---|---|---|
| Kriterium | Items pro Faktor | N | Items pro Faktor | N | Items pro Faktor | N |
| sehr gut | 4 | 500 | 4 | 900 | 4 | 1400 |
| | 6 | 250 | 6 | 200 | 6 | 260 |
| | 8 | 100 | 8 | 130 | 8 | 130 |
| gut | 5 | 130 | 5 | 140 | 5 | 200 |
| | 7 | 55 | 8 | 60 | 8 | 80 |

Vor der Durchführung der EFA sollte geprüft werden, ob sich die Korrelationsmatrix sowie die einzelnen Items zur Durchführung einer Faktorenanalyse eignen. Ausreißerwerte können sich verzerrend auf die Zusammenhangsmaße auswirken und sollten deshalb dringend ausgeschlossen werden (Bühner, 2011, S. 343). Die Korrelationsmatrix ist danach nach Variablen zu inspizieren, die mit keiner anderen Variable oder aber zu hoch ($r =| .90 |$) korrelieren. Betroffene Variablen müssen von den weiteren Analysen ausgeschlossen werden (Field, Miles & Field, 2012, S. 774). Zusätzlich sollte die Korrelationsmatrix mit dem *Kaiser-Mayer-Olkin-Kriterium* (KMO), der *measure sampling adequacy* (MSA) Prüfgröße sowie dem *Bartlett-Test* geprüft werden. Der Bartlett-Test überprüft dabei, ob die Korrelationsmatrix sich signifikant von null unterscheidet, d.h. ob überhaupt Zusammenhänge vorliegen, die auf latente Variablen zurückgeführt werden könnten. Wird der Bartlett-Test signifikant, unterscheiden sich die Korrelationen von null und die Korrelationsmatrix kann faktorisiert werden. Zu berücksichtigen ist allerdings, dass der Bartlett-Test von der Stichprobengröße abhängig ist und seine Signifikanz deshalb nur als Minimalbedingung betrachtet werden sollte (Bühner, 2011, S. 347 f.). Der KMO prüft, ob die in der Korrelationsmatrix vorliegenden Zusammenhänge groß genug sind, sodass diese sich zur Durchführung einer Faktorenanalyse eignet, während der MSA die Einzigartigkeit der Items prüft und demnach angibt, ob die einzelnen Items mit den übrigen Items hoch genug korrelieren (Bühner, 2011, S. 346 f.). Der Wertebereich des KMO und MSA liegen zwischen 0 und 1 (ebd.). Kaiser und Rice (1974) bieten eine Interpretationshilfe für die Bewertung der KMO- und der MSA-Maße, die in Tabelle 4.2 (S. 157) dargestellt ist.

Nach der Prüfung der Eignung der Korrelationsmatrix muss, auf Basis der zugrundeliegenden Verteilung der Daten, eine geeignete Extraktionsmethode gewählt werden (Baglin, 2014). Zu den bekanntesten und am häufigsten eingesetzten Extraktionsmethoden gehören die *Maximum-Likelihood* (ML) und die *Hauptachsenanalyse* (Principal Axis Factoring, PAF)

TABELLE 4.2: Eignung der Daten nach dem KMO- und MSA-Maß nach Kaiser und Rice (1974)

| KMO/MSA-Koeffizient | Eignung |
|---|---|
| $> .90$ | sehr gut (marvelous) |
| .80 - .90 | verdienstvoll (meritorious) |
| .70 - .80 | mittel (middling) |
| .60 - .70 | mäßig (mediocre) |
| .50 - .60 | schlecht (miserable) |
| $< .50$ | inakzeptabel (unnacceptable) |

(Bühner, 2011) neben Methoden wie *Unweighted Least Squares* (ULS) und *Generalized Least Squares* (GLS) (Costello & Osborne, 2005). Sowohl in der ML-Methode als auch der PAF werden die Zusammenhänge zwischen den Items auf eine geringe Anzahl von latenten Variablen zurückgeführt. Es wird also versucht die „Korrelation der Items durch wenige, homogene Faktoren vorherzusagen“ (Bühner, 2011, S. 320). In beiden Methoden wird die den Items gemeinsame Varianz untersucht (ebd.). Die ML-Methode wird empfohlen, wenn die Daten multivariat normalverteilt und intervallskaliert sind, die PAF-Methode dagegen, wenn dies nicht der Fall ist, da ihr keine Verteilungsannahmen zugrunde liegen (vgl. Baglin, 2014; Fabrigar et al., 1999; Schmitt, 2011). Zu den anderen Methoden fehlen diesbezüglich dagegen klare Richtlinien (Baglin, 2014).

Für die Schätzung der Faktorenanzahl können neben den inhaltlich-theoretischen Überlegungen auch deskriptivstatistische und inferenzstatistische Methoden zur Bestimmung der Faktorenanzahl herangezogen werden (Bühner, 2011). Zu den deskriptivstatistischen Methoden zur Bestimmung der adäquaten Faktorenanzahl gehören der *Scree-Test* nach Cattell (auch genannt Ellbogen-Kriterium) und das *Kaiser-Guttman-Kriterium* (Eigenwerte $> 1$). Die Verfahren *Acceleration Factor* (AF) und *Optimal Coordinate* (OC) von Raiche, Riopel und Blais (2006) und Raiche, Walls, Magis, Riopel und Blais (2013) stellen Weiterentwicklungen des Scree-Tests dar, die das Problem der subjektiv abhängigen Interpretation des Scree-Tests lösen (Raiche et al., 2013). Zu den inferenzstatistischen Methoden gehören beispielsweise die Parallelanalyse (PA, Horn-Kriterium Horn, 1965) und die MAP Methode (*Minimum Average Partial*, Velicer, 1976).

**Scree-Test:** Im Rahmen des Scree-Tests von Cattell (1966) werden die Eigenwerte nach abnehmenden Werten in ein Koordinatensystem eingetragen (Backhaus, Erichson, Plinke & Weiber, 2003, S. 296). Anhand dieses Eigenwertediagramms (auch Scree-Plot genannt) wird die bedeutsame Faktorenanzahl abgelesen. Die dargestellte Kurve zeigt die identifizierten Faktoren mit ihren jeweiligen Eigenwerten. Der Scree-Test wird auch Elbow-Kriterium genannt, da die Faktoren als bedeutsam angesehen werden, deren Eigenwerte vor dem Knick in der Kurve liegen (Bortz & Schuster, 2010, S. 415 f.).

**Kaiser-Guttman-Kriterium:** Nach dem Kaiser-Guttman-Kriterium (Guttman, 1954; Kaiser & Dickman, 1959) entspricht die Anzahl der bedeutsamen Faktoren der Anzahl der Faktoren

mit Eigenwerten über eins. Das führt jedoch häufig zu einer Überschätzung der Faktorenanzahl (Bortz & Schuster, 2010, S. 415 f.).

**Horn-Kriterium:** Nach diesem Kriterium nach (Horn, 1965) wird die Faktorenanzahl auf Basis einer Parallelanalyse abgeschätzt. Bei der Parallelanalyse wird der Eigenwerteverlauf der Korrelationsmatrix der erhobenen Daten mit dem Eigenwerteverlauf der Korrelation zwischen normalverteilten Zufallsvariablen verglichen. Ergebnis dieser Analyse ist eine grafische Darstellung, die in zwei Kurven beide Eigenwertverläufe zeigt. Die Anzahl der Faktoren, die vor dem Schnittpunkt der beiden Kurven liegen, entspricht der Anzahl der bedeutsamen Faktoren (Bortz & Schuster, 2010, S. 416).

**MAP:** „Beim MAP-Test werden so lange Komponenten aus der Korrelationsmatrix herauspartialisiert bis keine systematischen Zusammenhänge der Items in der verbleibenden Korrelationsmatrix mehr bestehen“ (Bühner, 2011, S. 328).

**AF:** Der AF ermittelt die Koordinate, an der sich das Gefälle der Kurve des Screeplots am stärksten ändert (Raiche et al., 2013, S. 24).

**OC:** Der OC ermittelt den Ort des Knicks im Screeplot, indem die Gradienten für alle Eigenwerte und die jeweils vorhergehenden Koordinaten aller anderen Eigenwerte ermittelt werden. Alle Eigenwerte werden dazu nacheinander untersucht. Es werden jeweils Linien von den Koordinaten des letzten Eigenwertes durch alle vorhergehenden Eigenwerte-Koordinaten gezogen. So kann für jeden beobachteten Eigenwert geprüft werden, ob er besser ist oder ungleich des geschätzten Eigenwertes (Raiche et al., 2013, S. 24).

Wie von einigen Autoren gezeigt werden konnte (Hubbard & Allen, 1987; Ruscio & Roche, 2012; Zwick & Velicer, 1986), neigen das Kaiser-Guttman-Kriterium und der Scree-Test jedoch dazu, die Anzahl der Dimensionen in den Daten zu überschätzen. Zwar ist der Scree-Test noch deutlich exakter als das Kaiser-Kriterium (Zwick & Velicer, 1986), allerdings kann das Ergebnis stark von der subjektiven Einschätzung der auswertenden Person geprägt sein (Baglin, 2014). Das OC- und AC-Verfahren haben sich in der Simulationsstudie von Ruscio und Roche (2012) gegenüber dem Kaiser-Kriterium als überlegen erwiesen. In den Simulationen von Ruscio und Roche (2012) konnte mittels der OC-Methode in 74% und der AF-Methode in rund 46% der Fälle die korrekte Faktorenanzahl ermittelt werden. Die OC-Methode erreichte damit fast so gute Ergebnisse wie die Parallelanalyse, die in rund 76% der Fälle die richtige Faktorenanzahl ermittelte. Die AF zeigte eine Tendenz zur Unterschätzung der Faktorenanzahl. Die Parallelanalyse und die MAP Methode stellen die geeignetsten Methoden zu Schätzung der Faktoranzahl dar, da diese in Simulationsstudien zuverlässig die korrekte Anzahl an zugrundeliegenden Dimensionen identifizieren konnten (Hubbard & Allen, 1987; Ruscio & Roche, 2012; Velicer, Eaton & Fava, 2000; Zwick & Velicer, 1986) und deshalb empfohlen werden (Bühner, 2011; Lance, Butts & Michels, 2006; Velicer et al., 2000).

Die inhaltliche Interpretation der Faktoren erfolgt auf Basis der Inhalte der zugehörigen Items. Um den Faktoren eine sinnvolle inhaltliche Benennung zu geben, sollten die Faktorladungen berücksichtigt werden. Sie zeigen, wie stark ein Faktor die jeweils gemessenen Items beeinflusst. Die Faktorstruktur muss untersucht werden, da sie zeigt, welche Items hoch auf welchem Faktor laden. Zur Interpretation der Faktoren wird geprüft, was die hoch auf einem Faktor ladenden Items inhaltlich gemeinsam haben. Eine Einfachstruktur, d.h. eine hohe Ladung der Items auf jeweils nur einem Faktor (ohne Nebenladungen), erleichtert dabei die Interpretation der Faktoren (Bühner, 2011). Ist eine Skala mehrdimensional, d.h. existieren mehrere Faktoren, muss zur Verbesserung der Interpretierbarkeit des resultierenden Faktormodells eine Faktorrotation vorgenommen werden (Beavers et al., 2013). Über die Anwendung eines Rotationsverfahrens kann die Interpretierbarkeit der Faktoren erhöht werden, indem möglichst eine Einfachstruktur hergestellt wird. Man unterscheidet dabei zwischen zwei Klassen, den *orthogonalen* und *obliquen Rotationsverfahren*. Der Einsatz ersterer ist angemessen, wenn keine Zusammenhänge zwischen den Faktoren bestehen. Die Varimax-Rotation stellt das bekannteste orthogonale Rotationsverfahren dar (ebd.). Ist jedoch von Zusammenhängen zwischen den Faktoren auszugehen, wie in den Sozialwissenschaften meist der Fall (Basto & Pereira, 2012, S. 3), führt ein obliques Rotationsverfahren zu genaueren Faktorlösungen (Baglin, 2014). Die zwei wichtigsten obliquen Verfahren sind Promax und Direct-Oblimin (Bühner, 2011, S. 338). Da vielfach gezeigt werden konnte (Costello & Osborne, 2005; Fabrigar et al., 1999), dass die oblique Rotation zu fast identischen Ergebnissen führt wie eine orthogonale Rotation, wenn die Faktoren wider Erwarten doch unkorreliert sind, sind oblique Verfahren insgesamt zu bevorzugen (Basto & Pereira, 2012, S. 3). Nach Beavers et al. (2013, S. 10) lässt sich bezüglich der obliquen Rotationsmethoden keine zu bevorzugende empfehlen. Welches der vielen obliquen Rotationsverfahren man wählt, hängt von der Verfügbarkeit in der Analysesoftware ab (Beavers et al., 2013; Fabrigar et al., 1999).

Bei der Interpretation der Faktoren nach der Rotation sollte darauf geachtet werden, dass resultierende Faktoren substanzielle Bedeutung haben. Sie sollten konzeptuell und empirisch relevant sein, insofern sie unter Rückgriff auf die Theorie inhaltlich interpretierbar sind (vgl. Bühner, 2011; Brown, 2015, S. 34). Schlecht definierte Faktoren sollten entfernt werden. Dabei handelt es sich um solche, auf denen eine zu geringe Anzahl von Items ($\leq$ 3 Items) lädt oder Faktoren, die nur aus Items mit geringen Kommunalitäten bestehen (vgl. Brown, 2015, S. 34). Die Interpretation der Faktoren erfolgt, wie oben bereits erwähnt, auf Basis der hoch ladenden Items. Ungeeignete Items sollten jedoch zunächst ausgeschlossen werden. Das betrifft Items mit starken Nebenladungen, d.h. solche, die nicht nur auf einem Faktor hoch laden (Hauptladung) sondern zusätzlich hohe Ladungen auf einem oder mehreren anderen Faktoren aufweisen (Nebenladungen). Auch Items mit geringen Ladungen auf allen Faktoren, also mit geringen Kommunalitäten, sollten entfernt werden (Brown, 2015, S. 34).

In einem letzten Arbeitsschritt muss das finale EFA-Modell, sofern im Rahmen der EFA Faktoren oder Items entfernt wurden, erneut an derselben Stichprobe mittels einer weiteren EFA überprüft werden (vgl. Brown, 2015, S. 34; Bühner, 2011). So lässt sich testen, ob die

Faktorstruktur auch nach der Item- und/oder Faktorreduktion Bestand hat (Bühner, 2011). Idealtypisch schließt sich bei einer Instrumententwicklung außerdem eine Prüfung der entstandenen Modells mittels CFA anhand einer neuen, größeren Stichprobe an (vgl. Brown, 2015, S. 34). Im Folgenden wird in die CFA eingeführt sowie ihre Teilschritte dargestellt.

### 4.3.2 Konfirmatorische Faktorenanalyse

Die konfirmatorische Faktorenanalyse (*confirmatory factor analysis*, CFA) ist eine Weiterentwicklung der EFA. Sie befasst sich speziell mit dem Messmodell, das definiert, welche manifesten Variablen die Indikatoren einer latenten Variable darstellen (Bühner, 2011, S. 391). Es geht also um die Beziehung der Items zu den jeweiligen Faktoren (Brown, 2015, S. 1). Die CFA erlaubt, die Ausprägungen in den Indikatorvariablen, d.h. die konkreten Itemantworten, kausal auf ein latentes Merkmal zurückzuführen. Die Itemantworten sind also Indikatoren für eine bestimmte Merkmalsausprägung und die Merkmalsausprägung, z.B. die vorherrschende epistemologische Überzeugung zur Quelle des Wissens, prädiziert ein bestimmtes Antwortverhalten, d.h. bestimmte Itemwerte (Bühner, 2011). Analysiert werden dazu Kovarianz- oder Korrelationsmatrizen. CFA, Pfadanalysen sowie die Kombination beider Verfahren werden häufig unter dem Begriff *lineare Strukturgleichungsmodelle* zusammengefasst. Pfadanalysen enthalten nur manifeste Variablen (Items), während die CFA sowohl latente als auch manifeste Variablen enthält (ebd.). Anders als bei der EFA werden alle Aspekte des Modells, das heißt die konkrete Anzahl der angenommenen Faktoren (latente Variablen) und die Zuordnung der Items zu diesen Faktoren (manifeste Variablen) vorher spezifiziert und danach überprüft (Brown, 2015, S. 1 f.). Es wird hier also ein vorher spezifiziertes Modell auf seine Passung zu den empirischen Daten überprüft, statt wie bei der EFA eine Faktorenstruktur zu ermitteln. Da die CFA in diesem Sinne a priori spezifizierte Hypothesen über die Struktur des Messmodells testet (Fabrigar et al., 1999), ist eine starke theoretische Basis in Bezug auf diese Strukturannahmen vonnöten (Bühner, 2011).

Abbildung 4.1 auf Seite 161 zeigt beispielhaft eine übliche grafische Darstellung so eines CFA-Modells mit zwei latenten Variablen $\xi$ und $\eta$, die als Kreise sowie insgesamt sechs manifeste Variablen $X_1$ bis $X_6$, die als Quadrate dargestellt sind. Doppelpfeile symbolisieren ungerichtete Zusammenhänge (Korrelationen oder Kovarianzen), während gerichtete Pfeile jeweils gerichtete Zusammenhänge darstellen. Die Stärke eines solchen gerichteten Zusammenhangs ist das Regressionsgewicht. In einer CFA wird dieses Regressionsgewicht als Faktorladung $\lambda$ bezeichnet. Jedes Item verfügt über eine Fehlervariable $\varepsilon$ (Itemfehlervarianzen), da jede Variable, die durch einen Prädiktor (latente Variable) vorhergesagt wird, über einen bestimmten Anteil an Varianz verfügt, der nicht vollständig vorhergesagt wird. Wird ein Parameter eines Modells nicht frei geschätzt sondern fixiert, wird der jeweilig festgelegte Wert eingetragen. Die Regressionsgewichte der Fehlervariablen sind hier (siehe Abb. 4.1, S. 161) jeweils auf 1 fixiert. Die latenten Variablen $\xi$ und $\eta$ beeinflussen jeweils die mit ihnen direkt verbundenen manifesten Variablen, allerdings wirken sie sich über ihre Korrelation $\varphi_{12}$ mit der jeweils anderen latenten Variable auch indirekt auf die manifesten Variablen

dieser latenten Variable aus. Die Größe der Korrelation $\varphi_{12}$ beeinflusst auch die Stärke dieses indirekten Einflusses. Bei der EFA wird im Unterschied dazu angenommen, dass alle manifesten Variablen direkt durch alle latenten Variablen beeinflusst werden.

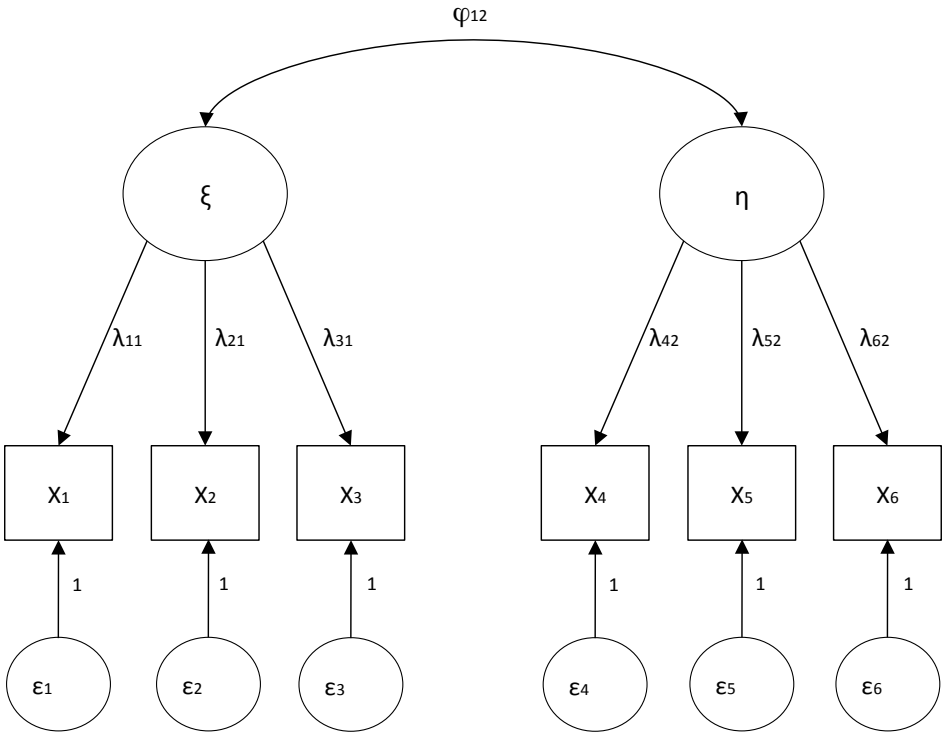


ABBILDUNG 4.1: Darstellung eines CFA-Modells mit zwei latenten Variablen ($\xi$ und $\eta$) und sechs Items ($X_1$ bis $X_6$)

Die Umsetzung der CFA gliedert sich in die folgenden Schritte (Bühner, 2011, vgl.):

- Prüfung der Voraussetzungen
- Spezifikation des Modells
- Wahl Schätzmethode
- Modelltest und Fit-Indizes
- Prüfung der Modifikationsindizes

Zunächst müssen die Voraussetzungen für die Durchführung einer CFA geprüft werden. Das betrifft die geforderte Anzahl der Items pro latenter Variable, die Stichprobengröße, Ausschluss von Ausreißern, Ausschluss von Kollinearität sowie Prüfung der Verteilung und Berücksichtigung des Skalenniveaus. In der CFA sollten mindestens vier manifeste Variablen pro latenter Variable spezifiziert werden (Bühner, 2011, S. 432), da mehr Items pro Faktor nach Marsh, Hau, Balla und Grayson (1998, S. 213) zu „angemesseneren Lösungen, exakteren und stabileren Parameterschätzungen, weniger nicht konvergierenden Lösungen

und reliableren Faktoren“ führen (Bühner, 2011, S. 432). Für die CFA wird eine größere Stichprobe als für die EFA benötigt, da kleine Stichproben von $N$ = 100 häufiger zu Problemen bei der Schätzung der benötigten Parameter führen (Marsh et al., 1998, S. 213). Je größer die Stichprobe für die CFA desto besser (Bühner, 2011). Bühner (2011, S. 432) empfehlen eine Stichprobengröße von mindestens 200 bis 250.

Ausreißer sowie Kollinearität sollten vor Durchführen der CFA ausgeschlossen werden. Ausreißer verzerren die Zusammenhänge, während sehr hohe Korrelationen zwischen Items (Kollinearität) zu Schätzproblemen führen. Bühner (2011, S. 432) schlägt als Richtlinie vor, dass Korrelationen zwischen Items r = .85 nicht überschreiten sollten.

Nach Prüfung der Voraussetzungen wird das Modell spezifiziert, das die Beziehungen zwischen den Items erklären können soll. Es werden die Beziehung der Items zu der jeweiligen latenten Variable (Messmodell) sowie die Beziehung, die zwischen den latenten Variablen untereinander besteht (Strukturmodell) definiert. Im Rahmen eines Modelltests prüft ein $\chi^2$-Test, ob die Modellstruktur die vorliegenden Daten der Stichprobe angemessen beschreibt. Dazu wird geprüft, ob die modelltheoretische Korrelations- oder Kovarianzmatrix des spezifizierten Modells zu der empirischen Korrelations- oder Kovarianzmatrix[2] der Stichprobe passt. Aus der empirischen Kovarianzmatrix werden dabei die unbekannten Parameter (Faktorladungen, Varianzen der latenten Variablen, Fehlervarianzen, Itemmittelwerte und Korrelationen) des Modells geschätzt. Dabei werden die Modellparameter so geschätzt, dass die Unterschiede zwischen der empirischen und der modellimplizierten Kovarianzmatrix möglichst gering ausfallen. Auf Basis der standardisierten Parameter erfolgt auch eine Signifikanzprüfung der jeweiligen Parameter. Welche Schätzmethode sich eignet, hängt vom Skalenniveau, der Art der Verteilung (normalverteilt oder schief) sowie der Stichprobengröße ab.

Die bekanntesten Schätzmethoden sind *Maximum-Likelihood* (ML), *Generalized Least Squares* (GLS), *Unweighted Least Squares* (ULS) und *Asymptotically Distribution-Free* (ADF). Welche Schätzmethode sich eignet, hängt vom Skalenniveau der Daten und der Verteilung ab. Die ADF-Methode, die auf einer GLS-Schätzung basiert, setzt keine Normalverteilung und Intervallskalenniveau voraus (Schermelleh-Engel, Moosbrugger & Müller, 2000; West, Finch & Curran, 1995). Nach Olsson, Foss, Troye und Howell (2000) führt die ADF-Methode nur bei einer Stichprobengröße ab $N$ = 1000 zu besseren Parameterschätzungen und einem genaueren Einschätzung der Modellpassung als die ML-Methode. Die ML-Methode erfordert, dass die manifesten Variablen aus einer multivariaten Normalverteilung stammen (Bühner, 2011, S. 409) und intervallskaliert sind (Bühner, 2011, S. 432). Liegt keine multivariate Normalverteilung vor, führt dies zu einer Überschätzung des $\chi^2$-Wertes des Modelltests und zu mittleren bis stärkeren Unterschätzungen der Standardfehler der Parameterschätzungen (Korrelationen und Ladungen) (Byrne, 2010, S. 268). Diese Standardfehler der Parameterschätzung werden benötigt, um die Signifikanz der Faktorladungen zu bestimmen. Insgesamt führt dies dazu, dass ein passendes Modell zu häufig abgelehnt wird (Bühner, 2011, S. 408) sowie Pfade und Korrelationen zwischen beobachteten Variablen, zwischen latenten Variablen oder zwischen beobachteten und latenten Variablen signifikant werden, obwohl dies

[2] Abhängig davon, ob die Kovarianz- oder Korrelationsmatrix analysiert wird.

für die Population nicht zutrifft (West et al., 1995, S. 73). Die ML-Methode ist die am häufigsten eingesetzte Schätzmethode der CFA und nach Bühner (2011, S. 413) die empfehlenswerte. Liegt eine Verletzung der multivariaten Normalverteilung vor, sollte beim Einsatz der ML-Methode der p-Wert für den Gesamttest mittels der Bollen-Stine-Bootstrap Methode korrigiert werden. Zu noch genaueren Parameterschätzungen und Teststatistiken unter Anpassung der $\chi^2$-Teststatistik führt allerdings der Einsatz der MLR (*Maximum-Likelihood Robust*), wenn die multivariate Normalverteilung nicht gegeben ist und Ratingskalen vorliegen (Schmitt, 2011).

Beim $\chi^2$-Modelltest wird die exakte Modellpassung (Modell-Fit) überprüft. Bei einem signifikanten $\chi^2$-Test gegen die Modellübereinstimmung (Fit) wird die exakte Modellpassung abgelehnt, das spezifizierte Modell konnte somit anhand der Daten nicht bestätigt werden. Hier ist allerdings zu berücksichtigen, dass der $\chi^2$-Test von der Stichprobengröße abhängig ist und größere Stichproben dafür sorgen, dass auch kleine Abweichungen von einem perfekten Modell-Fit bereits zu einer Ablehnung der Nullhypothese führen, da dass das Modell zu den empirischen Daten passt. Da für die Durchführung einer CFA meist größere Stichproben eingesetzt werden, u.a. weil die Schätzmethoden diese voraussetzen und bei kleinen Stichproben Schätzprobleme auftreten, ist eine Modellablehnung durch den $\chi^2$-Test nicht ungewöhnlich. Er wird deshalb häufig kritisiert, sollte dennoch immer angegeben werden, da er zeigt, ob exakter Modell-Fit vorliegt (Brown, 2015; Bühner, 2011). Aufgrund der Kritik wird zur Beurteilung der Modellpassung allerdings vorgeschlagen, das Ergebnis des $\chi^2$-Tests eher als deskriptiven Index für die Modellpassung zu nutzen. Für einen akzeptablen Fit sollte das Verhältnis des $\chi^2$-Wertes zu den Freiheitsgraden *df* gelten: $2 \leq \chi^2/df \leq 3$ und für gute Modellpassung $0 \leq \chi^2/df \leq 2$ (Schermelleh-Engel et al., 2000).

Zusätzlich werden zur Bestimmung der Modellpassung die Fit-Indizes (approximativer Modell-Fit) herangezogen (Brown, 2015). Die verschiedenen Fit-Indizes vergleichen das spezifizierte Modell immer mit einem alternativen Modell. Die inkrementellen oder auch *komparativen Fit-Indizes* nutzen als solches Vergleichsmodell ein sogenanntes *Nullmodell* (ebd.). Dabei handelt es sich um ein Modell, bei dem nur die Varianzen der Items geschätzt werden und alle anderen Parameter auf null fixiert sind, also davon ausgegangen wird, dass alle Variablen unkorreliert sind (R. B. Kline, 2016). Sie geben die proportionale Verbesserung der Anpassung, die das spezifizierte Modell im Vergleich zum Nullmodell erreicht, in einem Wertebereich von 0 bis 1 an. Höhere Werte sagen besseren Modell-Fit aus. Ein bedeutender Vertreter dieser Art von Fit-Indizes ist der *Bentler Comparative-Fit-Index* (CFI, Bentler, 1990). Bei den komparativen Fit-Indizes wird je nach Qualität des Index zwischen drei Typenklassen unterschieden (Typ-1-, Typ-2- und Typ-3-Index). Da bei der Berechnung des CFI viele Parameter berücksichtigt werden, handelt es sich um einen der besten komparativen Fit-Indizes von Typ-3 (Bühner, 2011, S. 424).

Die andere Art der Fit-Indizes nutzt als Vergleichsmodell eines, das exakt die in der Stichprobe vorhandenen Kovarianzen repliziert und die Daten perfekt beschreibt, ein sogenanntes *saturiertes Modell*. Diese sogenannten *absoluten Fit-Indizes* geben entweder an, wie gut oder wie schlecht das spezifizierte Modell die Daten im Vergleich zum Vergleichsmodell beschreibt (Bühner, 2011, S. 424 ff.).

Fit-Indizes, die angeben, wie gut die Modellpassung ist (*Goodness-of-Fit*), werden von solchen unterschieden, die angeben, wie schlecht die Modellpassung (*Badness-of-Fit*) ist. Bei den *Goodness-of-Fit-Indizes*, wie dem *Goodness-of-Fit-Index* (GFI), drückt ein möglichst hoher Wert, nahe 1, eine gute approximative Modellpassung aus, während bei den *Badness-of-Fit-Indizes*, wie dem *Steiger-Lind Root Mean Square Error of Approximation* (RMSEA, Steiger, 1990) oder der *Standardized Root Mean Square Residual* (SRMR), ein Wert möglichst nahe 0 angestrebt wird, der einen guten approximativen Modell-Fit aussagt.

Neben dem GFI, CFI, SRMR, TLI und RMSEA gibt es noch viele weitere Fit-Indizes, die allerdings aufgrund ihrer schlechten Performanz in Simulationsstudien (Hu & Bentler, 1998; Marsh, Balla & McDonald, 1988) nicht empfehlenswert sind (Brown, 2015). R. B. Kline (2016, S. 269) empfiehlt zusätzlich zum $\chi^2$-Wert sowie dem zugehörigen Freiheitsgrad und dem $p$-Wert, immer den CFI, RMSEA und den SRMR anzugeben, da diese unterschiedliche Informationen über die Modellpassung bereitstellen.

**RMSEA:** Beim RMSEA handelt es sich um einen absoluten Badness-of-Fit-Index, weshalb ein kleiner Wert nahe 0 eine gute Modellpassung aussagt. Er stellt die durchschnittliche Abweichung der Daten vom spezifizierten Modell je Freiheitsgrad dar und berücksichtigt sowohl den Stichprobenumfang als auch die Modellkomplexität. Als Richtlinie für den RMSEA gilt RMSEA < .60 bei $N > 250$ und RMSEA < .80 bei $N \leq 250$ (Hu & Bentler, 1999, vgl. Bühner, 2011). Nach Browne und Cudeck (1993) spricht ein RMSEA < .50 für guten Modell-Fit. Ein RMSEA von 0 deutet auf perfekten Modell-Fit hin. Für den RMSEA lässt sich das neunzigprozentige Konfidenzintervall (CI90) berechnen sowie der p-close-fit (CFit, Browne & Cudeck, 1993) angeben. Für das Konfidenzintervall gilt für akzeptablen bzw. guten Modell-Fit, dass die linke Grenze möglichst nahe 0 und die rechte Grenze unter dem Cut-off-Wert für den RMSEA liegen sollte (Hu & Bentler, 1999). Ist der linke Wert des Konfidenzintervalls 0 und der rechte Wert < .50 deutet dies wiederum auf perfekten Fit hin. Höhere Modellkomplexität führt zu breiterem Konfidenzintervall (Brown, 2015). Beim p-close-fit wird die Nullhypothese geprüft, dass der RMSEA unter .50 ist, d.h. der RMSEA eine gute Modellpassung bescheinigt. Wird die Hypothese *nicht* signifikant ($p \geq .10$), kann die Nullhypothese beibehalten werden und demnach angenommen werden, dass das spezifizierte Modell akzeptabel ($p \geq .10$) oder sogar gut ($p \geq .15$) zu den empirischen Daten passt (Brown, 2015, S. 274 f.).

**SRMR:** Auch der SRMR ist ein Badness-of-Fit-Index, weshalb ein kleiner Wert nahe 0 für einen guten Modell-Fit angestrebt wird. „Er kennzeichnet die standardisierte durchschnittliche Abweichung (Residuum) zwischen der beobachteten und der implizierten Korrelationsmatrix."(Bühner, 2011, S. 427). Er berücksichtigt zwar nicht die Modellkomplexität, allerdings verfügt er über eine geringe Sensitivität gegenüber dem Stichprobenumfang. Für den SRMR wird als Richtlinie für die Modellpassung teilweise SRMR um .08 oder kleiner bzw. SRMR < .11 angegeben (vgl. Bühner, 2011; Hu & Bentler, 1999).

**CFI:** Es handelt sich hierbei um einen komparativen Goodness-of-Fit Index, d.h. ein Wert nahe 1 signalisiert guten Modell-Fit im Vergleich zum Nullmodell. Der CFI vergleicht das spezifizierte Modell mit dem restriktiveren Nullmodell. Der CFI ist robust gegenüber Verletzungen der Verteilungsannahme und wenig sensitiv gegenüber der Stichprobengröße (Brown, 2015, S. 72). Nach Hu und Bentler (1999) sollte der CFI für eine Akzeptanz des Modells um .95 liegen, während Bentler (1990) einen akzeptablen Modell-Fit ab CFI = .90 sehen.

Für eine genauere Beschreibung der Fit-Indizes sei auf R. B. Kline (2016, S. 262-298) sowie Brown (2015, S. 67 ff.) verwiesen. In Tabelle 4.3 werden die wichtigsten Modell-Fits mit den oben genannten zugehörigen Beurteilungskriterien für gute Modellpassung und akzeptable Modellpassung zusammenfassend dargestellt.

TABELLE 4.3: Richtlinien für die Bewertung der Modellpassung nach den Fit-Indizes

| Fitindizes | exakter Modell-Fit | guter Modell-Fit | akzeptabler Modell-Fit |
|---|---|---|---|
| $\chi^2$ | p > .05 | | |
| $\chi^2/df$ | | $0 < \chi^2/df < 2$ | $2 < \chi^2/df < 3$ |
| RMSEA | 0 | < .05 | < .06 |
| RMSEA CI 90 | linke Grenze = 0, rechte Grenze < .05 | linke Grenze nahe 0, rechte Grenze < .05 | < .06 |
| RMSEA CFit | | .15 < p < 1.00 | .05 < p < .10 |
| SRMR | | < .08 | .08 < SRMR < .11 |
| CFI | | 0.95 | .90 < CFI < .95 |

Liegt der approximative Modell-Fit innerhalb der empfohlenen Grenzen, bei gleichzeitig signifikantem $\chi^2$-Test, muss anhand der Modifikationsindizes geprüft werden, wo Abweichungen im Modell vorliegen. Die Modifikationsindizes geben Hinweise für zu spezifizierende Pfade, die zu einer Modellverbesserung im Sinne eines Absenkens des $\chi^2$-Wertes führen würden. Es ist allerdings dringend darauf zu achten, dass sich solche zusätzlichen Spezifikation von Ladungen, kovariierenden Itemfehlervarianzen und Kovarianzen auch inhaltlich und theoretisch sinnvoll erlären lassen müssen, bevor sie umgesetzt werden (Bühner, 2011, S. 429).

Die CFA erlaubt es, konkrete Annahmen zu Pfaden zu testen, die Eindimensionalitätsannahme zu prüfen und über die Modifikationsindizes bietet sie Hinweise für eine Modellrevision an. Darüber hinaus lässt sich die Äquivalenz der Messmodelle prüfen. Äquivalente Messmodelle, wie das parallele, essentiell tau-äquivalente und tau-kogenerische (siehe Kap. 4, S. 145) können geprüft werden, indem jeweils die Bedingungen dieser verschiedenen Messmodelle im Sinne von Modellrestriktionen spezifiziert und deren Modellpassung getestet und untereinander verglichen wird (siehe Abb. 4.2, S. 166). „Werden die Modelle unter Berücksichtigung der entsprechenden Restriktionen nicht abgelehnt, sind die Items, Testteile bzw. Tests essentiell parallel, essentiell tau-äquivalent oder tau-kongenerisch." (Bühner, 2011, S. 401). So kann also identifiziert werden, welches Messmodell das Verhältnis zwischen manifesten und latenten Variablen am besten beschreibt und welches nicht zutrifft.

Dies stellt eine Voraussetzung für die Identifikation der geeigneten Reliabilitätsschätzung dar.

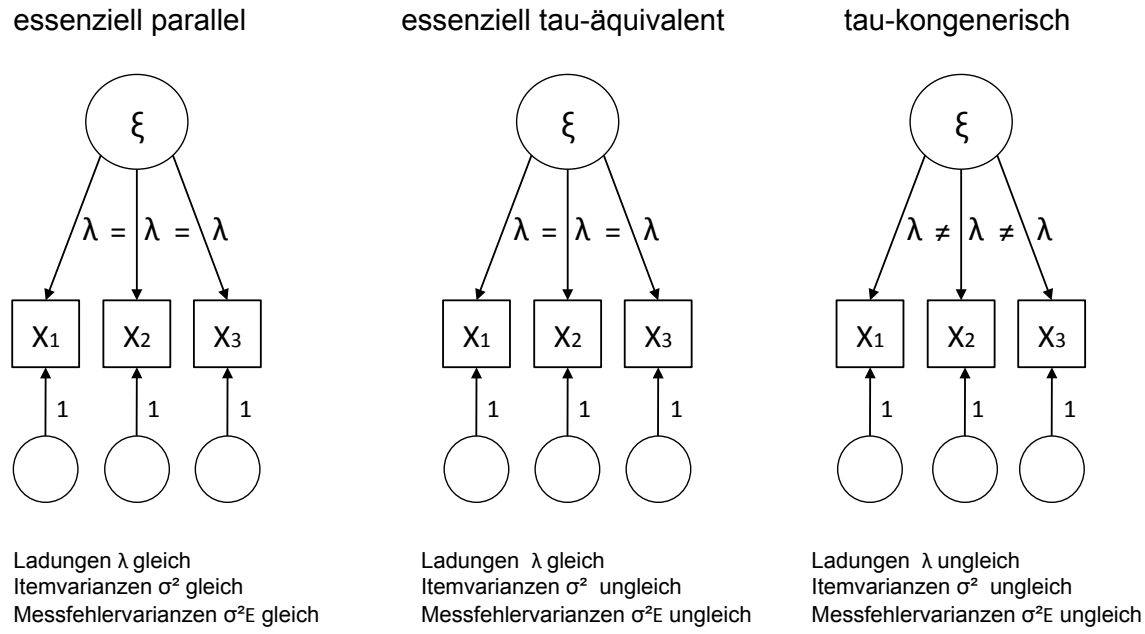


ABBILDUNG 4.2: Darstellung verschiedener Stufen der Äquivalenz von Messungen nach Bühner (2011, S. 400), in eigener Darstellung

Es lassen sich nach Jöreskog (1993) drei Arten des Einsatzes der CFA unterscheiden. Beim *strikt konfirmatorischen Ansatz* wird ein einziges Modell getestet, das entweder Modellpassung aufweist und als geeignet angenommen wird oder keine Modellpassung aufweist und demnach verworfen wird. Im Gegensatz dazu werden im Rahmen des *modellvergleichenden Ansatzes* anhand eines empirischen Datensatzes mehrere alternative Modelle miteinander verglichen und das geeignetste ausgewählt. Beim *modellgenerierenden Ansatz* wird ein tentatives Modell mittels eines empirischen Datensatzes auf Passung überprüft. Bei fehlender Passung wird dieses Modell angepasst und erneut getestet. Die Revision des Modells basiert dabei auf theoretischen Überlegungen oder erfolgt auf Grundlage der Daten. Der *modellgenerierende Ansatz*, d.h. der explorativere Einsatz der CFA (Bollen, 1989; Browne, 2001), ist im Rahmen dieser Arbeit der geeignete, da die CFA in dieser Art als ein Instrument zur Konstruktion und Evaluation von Skalen eingesetzt werden kann. Allerdings werden in dieser Instrumententwicklung die theoretischen Überlegungen, nach der Forderung von Bühner (2011, S. 429) bei der Modellrevision stets im Vordergrund stehen. Eine Modellrevision allein auf Grundlage der Daten wird nicht erfolgen.

Für die Testentwicklung nach der klassischen Testtheorie bietet es sich an, die EFA und den modellgenerierenden Ansatz der CFA zu kombinieren (Bühner, 2011). So kann in einem ersten Schritt und anhand einer ersten Stichprobe die Dimensionalität und Faktorstruktur

des Testentwurfes explorativ ermittelt werden und in einem zweiten Schritt anhand einer neuen, größeren Stichprobe geprüft werden.

Dieses kombinierte Vorgehen wird für diese Testentwicklung einerseits gewählt, weil es sich um das idealtypische Vorgehen der Testkonstruktion nach der KTT handelt (Brown, 2015; Bühner, 2011). Auf der anderen Seite eignet sich dieses Vorgehen gerade auch für das spezifische Zielkonstrukt dieser Testentwicklung sehr gut. Denn wie aus der Darstellung des theoretischen Hintergrunds epistemologischer Überzeugungen hervorging, existieren zahlreiche Konzeptionen epistemologischer Überzeugungen und die theoretischen Grenzen sind noch unscharf. Das ist auch daran zu erkennen, dass die Überzeugungen zu Lernen, die eigentlich nicht zum Kernkonstrukt epistemologischer Überzeugungen gehören, in vielen Skalen nach wie vor enthalten sind. Es handelt sich bei den Skalen epistemologischer Überzeugungen eher um tentative Modelle (Pintrich, 2002), wie schon daran zu erkennen ist, dass auch beim Einsatz bestehender und sehr häufig verwendeter Tests, wie dem SEQ, vielfach neue Faktorstrukturen identifiziert werden. Trotz eines starken theoretischen Hintergrundes (siehe Kapitel 2.10, ab S. 126) und starken Annahmen zur Dimensionalität (siehe Testentwurf Kapitel 5, ab S. 169), ist es folglich sinnvoll, anhand von zwei getrennten Stichproben zuerst die EFA und danach die CFA einzusetzen (vgl. Brown, 2015, S. 20, Bühner, 2011), statt beispielsweise gleich anhand einer ersten Stichprobe die CFA oder gar nur Trennschärfeanalysen einzusetzen.

# 5 Entwurf des StEB Inventars

Zielsetzung dieser Arbeit ist die methodisch saubere Konstruktion eines Tests zur mehrdimensionalen und disziplinspezifischen Erfassung der epistemologischen Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen aus Studientexten. Eine solche systematisch durchgeführte Konstruktion eines mehrdimensionalen Tests basiert auf einer „theoretisch begründeten, genau festgelegten Zahl inhaltlich klar umrissener Teilkomponenten (Faktoren) des Zielkonstrukts“ (Bortz & Döring, 2006, S. 221). Die angenommenen Dimensionen werden also stringent aus dem theoretischen Basismodell abgeleitet und als Subskalen des Tests operationalisiert, sodass für jeden Faktor später ein separater Testwert berechnet wird (vgl. Bortz & Döring, 2006, S. 221).

Die theoretische Basis, die den Anfang dieser Testentwicklung darstellt, wurde im Theorieteil dieser Arbeit umfassend dargestellt und im Kapitel 2.10 Synthese (S. 126) vereint. Hieraus geht hervor, dass das Modell von Hofer und Pintrich (1997) und Hofer (2000, 2001) die zentrale theoretische Basis für die Skalenkonstruktion und Itemformulierung darstellt. Die Kerndimensionen epistemologischer Überzeugungen sind Überzeugungen zur *Quelle des Wissens*, *Rechtfertigung des Wissen*, *Sicherheit des Wissens* und *Einfachheit des Wissens* nach Hofer und Pintrich (1997) und auf Conley et al. (2004) basiert die zusätzliche Dimension *Entwicklung des Wissens*. Die Überzeugungen zur Sicherheit, Einfachheit und Entwicklung des Wissens können dem übergeordneten Bereich *Beschaffenheit des Wissens* zugeordnet werden, während die Überzeugungen zur Quelle und Rechtfertigung des Wissens zum Bereich *Beschaffenheit des Wissensprozesses* gehören. Diese Dimensionierung stellt, der Forderung von Bortz und Döring (2006, S. 221) entsprechend, die Basis für die Entwicklung des Textentwurfs dar.

Die *Entwicklung des Testentwurfs* umfasst die folgenden zentralen Arbeitsschritte (Bühner, 2011, S. 84-107): Zu Beginn wird festgelegt, ob der Test *subjektive* oder aber *objektive Indikatoren* einer Eigenschaft oder Fähigkeit messen wird. Des weiteren ist die *Zielgruppe* zu bestimmen und das *Testziel* festzulegen. Das Testziel bestimmt, ob es sich bei dem zu entwickelnden Test um einen Wissenstest handeln soll, ob er auf die Trennung bzw. Klassifikation von Personengruppen oder aber die Bestimmung von Eigenschafts- oder Fähigkeitsausprägungen zielt. Vom Testziel ist abhängig, ob sich eher eine *rationale* (deduktive), *externale* oder *induktive Konstruktionsstrategie* eignet. Dem schließt sich, unter Einsatz von Top-Down- und/oder Bottom-Up-Ansätzen, das *Eingrenzen des Konstrukts*, *Generieren von Indikatoren* und die *Definition des Messgegenstandes* an. Abschließend wird das *Itemformat* gewählt und unter Berücksichtigung der *Formulierungsrichtlinien* ein Pool von Items formuliert (vgl. Bühner, 2011, S. 108 ff., S. 133 ff.). Diese Teilschritte, die die Entwicklung des Testentwurfs darstellen, werden in den folgenden beiden Kapitel beschrieben.

## 5.1 Dimensionen des StEB

Im Rahmen der *Konstrukteingrenzung* wird zunächst der Konstruktraum bestimmt, d.h. es wird festgestellt, was inhaltlich zum Konstrukt gehört, wo Grenzen liegen und aus welchem Verhalten, Eigenschaften und Einstellungen (Indikatoren) das Konstrukt besteht. Außerdem wird das Konstrukt in ein nomologischen Netzwerk eingeordnet, indem es im Verhältnis zu angrenzenden Konstrukten positioniert und Überlappungsbereiche sowie Verbindungen zu anderen Konstrukten festgestellt werden (vgl. Bühner, 2011, S. 97 ff.). Dies wurde im Rahmen des theoretischen Hintergrundes dieser Arbeit vorgenommen, indem top-down anhand von existierenden Konzeptualisierungen epistemologische Überzeugungen definiert und theoretisch in der Kognition und Metakognition positioniert sowie bottom-up anhand von empirischen Befunden die Verbindungen zum lernrelevanten Konstrukt Lernstrategien aufgezeigt wurden. Es wurde hierbei also eine Kombination aus Top-Down- und Bottom-Up-Techniken eingesetzt, um das Konstrukt einzugrenzen. Auch die *Definition des Messgegenstandes* wurde top-down-orientiert, basierend auf verschiedenen theoretischen Konzeptionen und unter Berücksichtigung zahlreicher empirischer Forschungsbefunde vorgenommen (vgl. Bühner, 2011, S. 105 ff.). Diese Definition wurde im Rahmen des Kapitels 2.10 Synthese festgehalten. Auf Basis der eindeutigen Definition und Abgrenzung des zu messenden Konstrukts können die eigentlichen Dimensionen und Skalen des Messinstruments abgeleitet werden. Dieser Instrumententwicklung liegt die *deduktive Methode* (auch genannt: *rationale Testkonstruktion*) als *Konstruktionsstrategie* zugrunde, da sie auf Basis einer gut ausgearbeiteten Theorie für das zu erfassende Konstrukt erfolgt (vgl. Bühner, 2011, S. 93). Das zugrundeliegende theoretische Modell gibt folglich vor, wie das Konstrukt definiert wird und welche inhaltlichen Teilaspekte dazu gehören.

Die Dimensionen des StEB Inventars, die nach der Itemformulierung den Subskalen entsprechen, wurden deduktiv und strikt aus den Beschreibungen der Dimensionen von Conley et al. (2004) und aus der Konzeptualisierung von Hofer und Pintrich (1997) und Hofer (2000, 2001) abgeleitet. Demnach wurden fünf Dimensionen definiert, welche die Hauptdimension *epistemologische Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten* in ihren Teilaspekten und so in ihrer ganzen Breite abdecken. Im folgenden werden diese fünf Dimensionen der Reihe nach definiert und dargestellt. Für jede der fünf Dimensionen wird aufgezeigt, welche Ausprägung epistemologischer Überzeugung einer wenig sophistizierten (Pol 1) und welche einer sophistizierten, fortgeschrittenen Merkmalsausprägung (Pol 2) entspricht. Außerdem wird für diese beiden Extreme explizit der disziplinspezifische Bezug zum Lernen mit erziehungswissenschaftlichen Studientexten hergestellt. Da das Modell von Hofer und Pintrich (1997) und dementsprechend deren Beschreibung der Kerndimensionen auf die Arbeiten von Schommer (1990), King und Kitchener (1994) und Kuhn (1991) zurückgeht, basieren auch deren Dimensionenbeschreibungen auf diesen Arbeiten. Aus diesem Grund beinhaltet die folgende Beschreibung der Dimensionen auch Verweise auf diese Autorinnen und Autoren.

**Quelle des Wissens**

Die Dimension *Quelle des Wissens* bezieht sich auf die epistemologische Vorstellung über den Ursprung des Wissens. Das betrifft die Überzeugung, dass Wissen entweder eher von außen (external) über Autoritäten an Lernende herangetragen wird oder eher auf der Fähigkeit der Lernenden aufbaut, Wissen in Interaktion mit anderen zu konstruieren (Hofer, 2000; Hofer & Pintrich, 1997).

**external:** Wissen wird vom Individuum als external – außerhalb des Selbst liegend – angesehen. Wissen liegt und hat seinen Ursprung in Autoritäten, die das Wissen an Personen übermitteln können.

**internal:** Das Individuum sieht in sich selbst den Ursprung von Wissen, in dem Sinne, dass in ihm/ihr selbst die Fähigkeit ruht, Wissen in Interaktion mit anderen zu konstruieren.

**Bezug zu Studientexten: external** Autoritäten sind in diesem Fall die Autorinnen und Autoren der Studientexte, die ihr Wissen in diesen Texten festhalten. Die Studientexte dienen dazu, das Wissen an die Studierenden zu übertragen, wenn diese die Studientexte lesen.

**Bezug zu Studientexten: internal** Wissen kann einerseits selbst durch Auseinandersetzung mit dem Studientext und unter Berücksichtigung weiterer Texte konstruiert werden oder aber in Auseinandersetzung mit anderen, z.B. Kommilitonen.

**Rechtfertigung des Wissens**

Die Dimension *Rechtfertigung des Wissens* umfasst, wie Lernende Behauptungen bewerten, von Belegen Gebrauch machen und wie sie Autorität und Expertise annehmen sowie Expertinnen und Experten bewerten. Individuen auf niedrigerem Level rechtfertigen Ansichten über Beobachtung, Autorität oder auf der Basis davon, was sich richtig anfühlt. Individuen auf höherer Stufe prüfen Behauptungen und wägen die Ansichten von Experten ab und integrieren sie (Hofer, 2000; King & Kitchener, 1994).

**absolut:** Rechtfertigung des Wissens erfolgt allein über Autorität oder Beobachtung. Richtig ist für das Individuum, was man beobachten kann, was Autoritäten annehmen oder – im Falle von Unsicherheit – was sich richtig anfühlt. Keine richtige Begründung für Meinungen und Positionen ist nötig und wird gegeben.

**evaluativ:** Wissen wird über Nachforschung und Prüfung begründet und gerechtfertigt. Das Individuum bewertet und integriert die Ansichten von Experten und neigt zu vielfältigen Meinungen und begründeten Urteilen.

**Bezug zu Studientexten: absolut** Richtig ist, was man mal gelesen hat bzw. was die Autorinnen und Autoren berichten. Die Glaubwürdigkeit ist abhängig von Stellung, Position und Titel des Autors oder der Autorin.

**Bezug zu Studientexten: evaluativ** Eine Meinung zum Studientext wird gebildet auf Basis kritischen Reflektierens der Textinhalte, genannter Forschungsergebnisse und im Abgleich mit dem Vorwissen und mit Positionen und Forschungsergebnissen anderer Autorinnen und Autoren. Es wird darauf Wert gelegt, dass die Autorinnen und Autoren logisch begründen, argumentieren und Gegenpositionen abwägen.

**Sicherheit des Wissens**

Die Dimension *Sicherheit des Wissens* umfasst die beiden epistemologischen Vorstellungen, dass Wissen einerseits eher als feststehend angenommen wird, in dem Sinne, dass es „die eine richtige Antwort" (Wahrheit) gibt. Auf der anderen Seite steht die Überzeugung, dass es keine absoluten Wahrheiten gibt und Wissen mit Unsicherheit behaftet ist. Individuen auf diesem Level gehen davon aus, dass komplexe Probleme mehrere richtige Antworten haben können. Sie sind neuen Interpretationen gegenüber offen und auch der Möglichkeit, die eigene Ansicht im echten Austausch mit anderen zu verändern (Conley et al., 2004; Hofer, 2000; King & Kitchener, 1994; Kuhn, 1991).

**absolut:** Wissen wird vom Individuum als feststehend wahrgenommen, d.h. es wird angenommen, dass es auf Fragen/Probleme ausschließlich eine richtige Antwort gibt.

**multiplizistisch:** Für das Individuum ist Wissen mit Unsicherheit behaftet und es gibt keine absoluten Wahrheiten. Es kann mehr als eine richtige Antwort auf Fragen oder Probleme geben. Neuen Interpretationen wird offen gegenüber gestanden, wie auch der Möglichkeit, die eigene Ansicht zu ändern.

**Bezug zu Studientexten: absolut** Was im Studientext steht ist entweder wahr oder falsch. Es gibt nur eine richtige Antwort auf eine Frage bzw. Ansicht zu einem Sachverhalt.

**Bezug zu Studientexten: multiplizistisch** Mehrere im Studientext berichtete Positionen könnten zutreffen oder nicht zutreffen. Was richtig ist, kann auch vom Kontext abhängig sein. Manche Ansichten treffen vielleicht wahrscheinlicher zu als andere. Es gibt keine absolute Wahrheit, deshalb ist der Austausch von Ansichten wichtig.

**Entwicklung des Wissens**

Die Dimension *Entwicklung des Wissens* umfasst die epistemologischen Überzeugungen zur Veränderlichkeit des Fachbereichs (d.h. z.B. Pädagogik als Disziplin). Auf der einen Seite steht die Auffassung, dass die Fachwissenschaft als solche nicht veränderlich ist, d.h. dass sich nicht ändert, was jetzt als wahr angenommen wird. Auf der anderen Seite wird die Disziplin als sich entwickelnd und veränderlich wahrgenommen. Auf der Basis neuer Entdeckungen, Studien und Forschungsergebnisse ändern Fachexpertinnen und -experten ihre Ansichten darüber, was in ihrem Fach zutrifft (Conley et al., 2004).

**statisch:** Das Individuum ist überzeugt, dass die Disziplin sich nicht über die Zeit verändert, sondern statisch ist. Was Fachexpertinnen und -experten heute als wahr annehmen, wird auch in Zukunft als wahr angenommen werden.

**entwickelnd:** Die Disziplin wird als über die Zeit veränderlich wahrgenommen. Neue Entdeckungen, neue Forschung und Forschungsergebnisse führen dazu, dass Fachexpertinnen und -experten ihre Ansichten darüber ändern, was zutrifft.

**Bezug zu Studientexten: statisch** Die in Studientexten vertretenen Ansichten verändern sich nicht über die Zeit, d.h. neuere und ältere Texte vertreten dieselben Positionen. Autorinnen und Autoren bleiben bei ihren Positionen.

**Bezug zu Studientexten: entwickelnd** Die in den Studientexten vertretenen Ansichten über das Fach und darüber, was im Fach zutrifft, können sich über die Zeit aufgrund von Forschung verändern. Die Ansichten und Positionen von Autorinnen und Autoren können sich ändern.

**Struktur des Wissens**

Die Dimension *Struktur des Wissens* betrifft die epistemologische Vorstellung über die Struktur von Wissen. Wissen wird auf einem Kontinuum betrachtet, das von Wissen als Anhäufung von Fakten bis hin zu Wissen als in gegenseitiger Beziehung stehende Konzepte reicht. Auf niedrigerem Level wird Wissen als diskrete (klar abgrenzbare), konkrete, zu wissende Fakten betrachtet, d.h. Wissen ist die Anhäufung von Fakten. Auf höherer Ebene wird Wissen als hoch vernetzte, komplexe Konzepte betrachtet (Schommer, 1990, 1994).

**isoliert:** Wissen wird vom Individuum als diskrete (klar abgrenzbare), konkrete, zu wissende Fakten betrachtet. Wissen stellt eine Anhäufung von Fakten dar.

**integriert:** Wissen wird als hoch vernetzte, komplexe Konzepte wahrgenommen. Wissen ist demnach nicht einfach sondern komplex.

**Bezug zu Studientexten: isoliert** Die Fakten in einem Studientext stellen das Wissen dar, das dieser vermitteln kann.

**Bezug zu Studientexten: integriert** Die Fakten, die ein Studientext enthält, sind nicht auch gleichzeitig das Wissen, das dieser enthält. Da Wissen komplex und hoch vernetzt ist, können Fakten nur unvollständiges Wissen sein. Verschiedene Positionen zu einem Thema müssen berücksichtigt werden, um etwas wissen zu können.

Nachdem zuvor aufgezeigt wurde, wie die Dimensionen definiert werden, schließt sich nun die Itemformulierung zu diesen fünf angenommenen Dimensionen an, mit der die konkrete Skalenkonstruktion umgesetzt wird.

## 5.2 Itemformulierung

Das Ziel des StEB Inventars soll es sein, *Eigenschaftsausprägungen* von Personen bezüglich der individuellen epistemologischen Überzeugungen zum erziehungswissenschaftlichen Wissen zu ermitteln. Er wird dementsprechend nicht dazu dienen, Personen zu klassifizieren (Gruppentrennung) oder Wissen zu erfassen. Aus dieser Zielsetzung ergibt sich, dass bei der Itemformulierung darauf geachtet werden muss, dass Items zur Erfassung des Konstrukts gefunden werden, die nur ein einziges Konstrukt erfassen, aber dieses in seinen unterschiedlichen Teilaspekten in der vollen Breite abdecken (vgl. Bühner, 2011, S. 93). Es handelt sich hierbei um sogenannte reflektive Indikatoren, die untereinander korrelieren und alle die Ausprägung der latenten Variable – epistemologische Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten – widerspiegeln und deren Zusammenhang demnach durch die Merkmalsausprägung erklärt werden kann (vgl. Bühner, 2011, S. 92). In dem zu entwickelnden StEB Inventar werden *subjektive Indikatoren*, im Sinne von Selbsteinschätzungen, dazu eingesetzt, das Zielkonstrukt epistemologische Überzeugungen zu erfassen (vgl. Bühner, 2011, S. 85). Die *Zielgruppe* des StEB Inventars sind deutschsprachige Lehramtsstudierende ab dem ersten Fachsemester und einem Alter von mindestens 17 Jahren (bei zwölfjährigem Abitur), aber eher ab 18 Jahren. Die fachliche Ausrichtung und Fächerkombination oder angestrebte Schulart der Zielgruppe ist jedoch nicht festgelegt. Es wird für die Items des StEB Inventars das geschlossene Itemformat *Ratingskala* gewählt, bei dem festgelegte Antwortmöglichkeiten in Abschnitten auf einem Merkmalkontinuum vorgegeben werden und die quantitative, subjektive Beurteilung einer Merkmalsausprägung durch die befragten Personen möglich ist (Bortz & Döring, 2006, S. 177). Da die Intervalle des Merkmalskontinuums von den befragten Personen als gleich groß bewertet werden sollen, geht man davon aus, dass die Stufen der Ratingskala als Intervallskala interpretiert werden können (vgl. Bortz & Döring, 2006, S. 177, 181 f.). Für die StEB-Rating-Items wird eine bipolare Antwortskala mit fünf Skalenstufen gewählt. Das heißt, die beiden äußersten Antwortmöglichkeiten stellen entgegengesetzte Pole dar (Bortz & Döring, 2006, S. 180 f.). Die fünf Antwortstufen werden sowohl verbal als auch numerisch beschriftet, um das Bearbeiten der Items zu erleichern (Bühner, 2011, S. 116). Das gewählte Antwortformat orientiert sich bezüglich der verbalen Beschriftung der fünf Stufen am fünfstufigen Antwortformat für Bewertungen nach Rohrmann (1978) (*trifft gar nicht zu*, *trifft wenig zu*, *trifft teils-teils zu*, *trifft ziemlich zu*, *trifft völlig zu*), damit der Bedeutungsgehalt der sprachlichen Skalenbeschriftung möglichst gut mit der Zahlenkodierung korrespondiert (siehe Bühner, 2011, S. 113).

Die Ratingskala ist aufgrund des geschlossenen Antwortformats einerseits ökonomisch in der Durchführung und Auswertung und andererseits auch auswertungsobjektiv (Bortz & Döring, 2006, S. 213 ff.). Aus diesen Gründen und zusätzlich auch, weil häufig davon ausgegangen wird, dass die resultierenden Daten als intervallskaliert behandelt werden können und sich demnach für eine Vielzahl von statistischen Analysen eignen, handelt es sich bei der Ratingskala um ein sehr etabliertes Itemformat (Döring & Bortz, 2016, S. 408). So setzen auch alle in Kapitel 2.9.2 (ab S. 114) dargestellten quantitativen Erhebungsinstrumente epistemologischer Überzeugungen Ratingskalen ein. Dieses Itemformat wird aufgrund seiner

genannten Vorteile auch für die Entwicklung des StEB Inventars gewählt. Das erhöht außerdem die Vergleichbarkeit des StEB Inventars mit den übrigen Fragebögen.

An Items, die gemeinsam eine Dimension bilden, werden nach Bortz und Döring (2006, S. 222) einige Minimalbedingungen gestellt, um als Testskala bezeichnen zu können:

1. Die Messung soll eindimensional sein. Die Items sollen das interessierende Merkmal einheitlich messen, also möglichst homogen sein.
2. Es sollten Items von unterschiedlicher Schwierigkeit entwickelt werden, d.h. eine hohe Streuung der Schwierigkeitsindizes vorliegen, damit sie möglichst viele Ausprägungsgrade des Merkmals repräsentieren.
3. Die Items sollten hohe Trennschärfe aufweisen, damit sie eindeutig Personen mit starker Merkmalsausprägung von Personen mit schwächerer Merkmalsausprägung trennen können.
4. Eine hohe Testobjektivität muss gesichert werden, indem die Vorschriften für die Auswertung der Itemantworten möglichst eindeutig formuliert sind.
5. Die Anzahl und Formulierung der Items sollten eine möglichst reliable Merkmalsmessung gewährleisten.
6. Die Validität der einzelnen Items und des Gesamttestwerts muss theoretisch begründet und empirisch belegt sein.

Aus Anforderung 1 ergibt sich, dass jedes Item nur eine Aussage enthalten sollte, um sicherzustellen, dass es sich nicht auf mehrere latente Variablen bezieht. Es ist sinnvoll, die Items noch vor dem ersten Einsatz bezüglich ihres Inhalts zu prüfen (vgl. Validierung ab S. 150). Dies kann durch Expertenurteile unterstützt werden, die Hinweise auf möglicherweise konfundierte Items geben können. Anforderung 1 betrifft allerdings gleichzeitig auch die Frage nach der Repräsentativität des Items für das Konstrukt, denn wenn die resultierende Skala nicht eindimensional ist, misst sie neben dem Zielkonstrukt noch andere Konstrukte (vgl. Validierung ab S. 150). Anforderung 6 betrifft ebenfalls das Verhältnis der Items zum Konstrukt (ehemals Konstruktvalidierung), da nach Einsatz des Tests zu überprüfen ist, ob die Items und die Dimensionen des Tests auch tatsächlich das Zielkonstrukt repräsentieren. Auch die Itemschwierigkeit (Anforderung 2), Trennschärfe (Anforderung 3) und Reliabilität (Anforderung 5) lassen sich nur durch den Einsatz des Tests prüfen, indem eine Item- und Skalenanalyse durchgeführt wird, während Anforderung 4 über das Itemformat Ratingskala sehr einfach sicherzustellen ist.

Über die Eingrenzung des Konstrukts konnten aus der sehr vielseitigen Theorie und den bestehenden Definitionen epistemologischer Überzeugungen spezifische für diese Instrumententwicklung relevante Konzeptualisierungen abgeleitet werden (vgl. Kapitel 2.10, S. 126). Aus diesen ausgewählten Konzeptualisierungen sind die folgenden quantitativen

Erhebungsinstrumente hervorgegangen, die für die Itemformulierung gesichtet wurden, da sie wichtige Hinweise für geeignete Indikatoren des Konstrukts bieten (vgl. Bühner, 2011, S. 97 ff.):

- *Schommer Epistemological Questionnaire* (SEQ, siehe Schommer, 1998, S. 561 ff für eine vollständige Auflistung der Items),
- *Disciplin-Focused Epistemological Belief Questionnaire* (DEBQ, siehe Hofer, 2000, S. 390 ff. für eine vollständige Auflistung der Items),
- *Epistemic Belief Inventory* (EBI, siehe Schraw et al., 2002, S. 275 für eine vollständige Auflistung der Items),
- das Instrument von Conley und Kollegen (siehe Conley et al., 2004, S. 202 f. für eine vollständige Auflistung der Items) sowie
- die von Urhahne und Hopf (siehe Urhahne & Hopf, 2004, S. 84 f. für eine vollständige Auflistung der Items) übersetzte Version des Instruments von Conley et al. (2004, S. 202 f.).

Nach dem Vorbild der bestehenden Instrumente wurden die Items, wie es typisch für Ratingskalen ist, in Form von Aussagen formuliert, bezüglich derer Befragte den Grad ihrer Zustimmung angeben können. Insgesamt wurde ein Itempool von 126 Items formuliert, die a priori den fünf oben beschriebenen Dimensionen zugeordnet wurden. Für die Dimension *Quelle des Wissens* wurden 31 Items entwickelt, für *Rechtfertigung des Wissens* 34, *Sicherheit des Wissens* 18, *Struktur des Wissens* 26 und für die Dimension *Entwicklung des Wissens* wurden 17 Items formuliert. 26 der ingesamt 126 entwickelten Items basieren dabei grundlegend auf bereits bestehenden Items aus den oben genannten Instrumenten. Aufgrund der Disziplinspezifität und der Ausrichtung auf die Studientexte sind jedoch auch die Items, die auf bestehenden Instrumenten basieren, stark abgewandelt. Tabelle 5.1 zeigt die Anzahl der entwickelten Items je Subskala sowie die Anzahl der davon invertiert gepolten Items und die Anzahl der Items je Subskala, die auf den oben genannten Instrumenten basieren.

TABELLE 5.1: Item-Skalen-Zuordnung des 126-Item StEB Inventar Testentwurfs

| Überzeugungs-dimensionen | Skalen-code | Item-anzahl | Itemanzahl invertiert | Itemanzahl abgewandelt |
|---|---|---|---|---|
| Quelle des Wissens | QW | 31 | 24 | 9 |
| Rechtfertigung des Wissens | RW | 34 | 16 | 4 |
| Sicherheit des Wissens | GW | 18 | 14 | 3 |
| Struktur des Wissens | SW | 26 | 16 | 4 |
| Entwicklung des Wissens | EW | 17 | 6 | 6 |
| Gesamt | | 126 | 76 | 26 |

In Tabelle 5.2 wird außerdem dargestellt, welche epistemologischen Überzeugungsausprägungen den jeweiligen Polen der Subskalen zugeordnet sind. Im Sinne der Formulierungshinweise für Items nach Döring und Bortz (2016, S. 409 f.) wurde bezüglich der Instruktion und der Items darauf geachtet, dass sich der Sprachgebrauch an der Zielgruppe *Lehramtsstudierender* ausrichtet. Die Instruktion wurde zudem möglichst klar und informativ gehalten und die Items möglichst präzise und leicht verständlich formuliert, wobei auf doppelte Verneinungen, Abkürzungen, Fremdworte, mehrdeutige Begriffe und ungewöhnliche Satzkonstruktionen verzichtet wurde (ebd.).

TABELLE 5.2: Ausprägungen epistemologischer Überzeugungen je Skala für die Pole des 126-Item StEB Inventar Testentwurfs

| Skalen-code | wenig sophistiziert linker Pol | sophistiziert rechter Pol |
|---|---|---|
| QW | external | internal |
| RW | absolut | evaluativ |
| GW | absolut | relativistisch/ evaluativ |
| SW | isoliert | integriert |
| EW | absolut/ statisch | entwickelnd |

Nach den Formulierungsrichtlinien wurde auch versucht, die Sätze möglichst kurz zu halten (ebd.), aber gleichzeitig eine gendergerechte Formulierung umzusetzen. Wichtige Wörter wurden sparsam hervorgehoben (für die Formulierungshinweise vgl. Döring & Bortz, 2016, S. 409 und auch Bühner, 2011, S. 133-139). Vor den Items, jeweils am Anfang der Seite, wurde die folgende Instruktion eingefügt: „Wie stark stimmen Sie den Aussagen zu in Bezug auf Ihren Studienbereich Pädagogik/ Schulpädagogik/Erziehungswissenschaft?“ Dies soll zusätzlich zur konkreten Formulierung der Items sicherstellen, dass die Aussagen ausschließlich auf die erziehungswissenschaftlichen Studienanteile bezogen werden (siehe Anhang A, S. 359). Die Disziplinspezifität wird im Rahmen der Items umgesetzt, indem in jeder Aussage explizit der Bezug zu diesen pädagogischen *Studientexten* hergestellt wird (Itembeispiel: gw_2 „Forschende sollten sich erstmal einig über ihre Aussagen werden, bevor sie diese in Studientexten anführen.“). Da jedoch nicht davon auszugehen ist, dass der Begriff *pädagogische Studientexte* von allen Personen der Zielgruppe gleich interpretiert wird, wurde dieser im Rahmen der Instruktion vor den Items explizit erklärt (siehe Anhang A, S. 359). Im Anhang B in den Tabellen B.1 bis B.7 (ab S. 369) ist der genaue Itemwortlaut einzusehen. Hier geht auch die a priori Zuteilung der Items zu den Subskalen sowie die Formulierungsrichtung der Items hervor. Es wird darüber hinaus aufgezeigt, welche Items auf welchen bestehenden Instrumenten basieren.

Bevor diese Items jedoch einer ersten Stichprobe von Lehramtsstudierenden vorgelegt werden können, sollte im Sinne der oben beschriebenen (siehe Kapitel 5.2, ab S. 174) Bedingungen für Testskalen von Bortz und Döring (2006, S. 222) nach Anforderung 1 der Inhalt

der Items geprüft werden. Über diese Prüfung kann gleichzeitig im Sinne des argumentationsbasierten Validierungsansatzes Evidenz für die Grundannahmen 1 und 2 (siehe S. 139) gesammelt werden. Die Expertenbefragung, die zu diesem Zweck durchgeführt wurde, wird im Folgenden beschrieben.

## 5.3 Expertenbefragung

Der Validierungsprozess kann zu Beginn einer Instrumententwicklung über Expertenurteile gestützt werden, die im Sinne des alten Validierungsverständnisses zur Inhaltsvalidität und Konstruktvalidität beitragen (vgl. Döring & Bortz, 2016, S. 445 f.) oder im Sinne des neues Verständnisses als empirische Evidenzen zur Stützung der Testwertinterpretation 1 (siehe S. 137) beitragen (AERA et al., 2014; Hartig et al., im Druck). Das eingesetzte Verfahren ähnelt dabei dem Befragungsverfahren der Delphi-Methode, die darauf abzielt, die Expertise von verschiedenen Sachverständigen in einem bestimmten Fachbereich für die Optimierung einer Problemlösung zu nutzen (Bortz & Döring, 2006, S. 261). Im vorliegenden Fall zielt die Expertenbefragung auf die Optimierung der Itemauswahl, des Iteminhalts und -formulierung sowie der internen Struktur des Instruments, indem die Zuordnung der Items zu den Dimensionen geprüft wird.

Im Zeitraum von August 2015 bis November 2015 wurde die Expertise von 10 Expertinnen und Experten aus den Bereichen Pädagogische Psychologie, Erziehungswissenschaft, Psychologie und Pädagogik von fünf deutschen Universitäten und des ZPID-Leibniz Instituts in Trier genutzt, um die Itemauswahl, -inhalt und -formulierung zu optimieren, die interne Teststruktur zu prüfen und so die Testwertinterpretationen 1 (siehe S. 137) zu stärken. Vier dieser Experten waren zu dem Zeitpunkt der Befragung habilitiert, vier promoviert und zwei diplomiert. Bei acht dieser Personen ist davon auszugehen, dass sie über Expertise im Bereich der epistemologischen Überzeugungen verfügen, da sie in diesem Bereich aktiv forschen und publizieren bzw. im Falle einer Person früher in diesem Bereich geforscht und publiziert haben. Drei dieser Expertinnen und Experten haben selbst schon Instrumententwicklung im Bereich der epistemologischen Überzeugungen betrieben. Die beiden Experten, die nicht im Bereich der epistemologischen Überzeugungen forschen, wurden mit in die Befragung aufgenommen, da sie ihren Forschungsschwerpunkte in der Einstellungsmessung bzw. in den empirischen Forschungsmethoden und der Instrumententwicklung haben. Von zwei Personen, die zugesagt hatten, ging kein Rücklauf ein. Eine weitere Personen hat eine Teilnahme abgelehnt, da ihre Forschung zu epistemologischen Überzeugungen ihrer Meinung nach zu lange zurück liegt. Eine weitere Person hat aus Zeitgründen nicht teilgenommen.

Die Erstansprache zur Akquise erfolgte zunächst telefonisch. Bei Bereitschaft zur Teilnahme wurde den Expertinnen und Experten per Email das Material zugesendet, das ein Anschreiben, eine Beschreibung der fünf Dimensionen sowie die Items enthielt. Die insgesamt 126 entwickelten Items wurden für die Expertenbefragung zunächst nach Skalen sortiert und mit einer elfstufigen Ratingskala versehen. Die Expertinnen und Experten wurden gebeten, ihre Einschätzung über die Eignung der Items in Bezug auf ihre jeweilig intendierte Skala anzugeben. Die Einschätzung erfolgte auf einer elfstufigen Thurstone-Skala (Thurstone & Chave, 1929) von 1 *Item unpassend* bis 11 *Item passend*. Darüber hinaus wurden die Expertinnen und Experten gebeten, mögliche Hinweise und Kritik zu den einzelnen

Items zu dokumentieren. Die Version des Befragungsinstruments, die den Experten vorgelegt wurde, lässt sich im Anhang A ab Seite 341 einsehen.

Als Maß für die zusammengefasste Expertenmeinung wurden die Itemmediane und Itemmittelwerte berücksichtigt. Nach der Empfehlung von (Häder, 2014, S. 184) wurden Quartilsmaße und Standardabweichung als Informationsquellen über die Streuung der Expertenmeinungen ebenfalls analysiert. Die Einschätzungen der Expertinnen und Experten wurden mittels eines Mediansplits ausgewertet. Liegt der Median eines Items unterhalb des Subskalenmedians, wird angenommen, dass das entsprechende Item nicht geeignet ist, die epistemologischen Überzeugungen zu messen, die gemessen werden soll. Es wurde demnach ermittelt, welche Items wegfallen würden, wenn der über die Expertenurteile hinweg ermittelte Median jedes Items in Bezug zum entsprechenden Subskalenmedian gesetzt wird und nur solche Items beibehalten werden würden, deren Median größer oder gleich des Subskalenmedians ist. Dies wurde einmal für alle zehn Expertinnen und Experten berechnet sowie für eine Auswahl von acht und sieben Expertinnen und Experten. Die Auswahl der acht Expertinnen und Experten stellt die Urteile derjenigen Personen dar, die tatsächlich Expertise im Forschungsbereich epistemologische Überzeugungen aufweisen. Ein Experte beurteilte die Items auffällig abweichend zu den übrigen sieben Expertinnen und Experten dieses Forschungsfeldes sowie zu den beiden Experten der Einstellungsmessung bzw. Instrumentenentwicklung. Deshalb wurde geprüft, inwiefern sich die durchschnittlichen Expertenurteile der Personen mit Expertise im Bereich epistemologische Überzeugungen ändern, wenn die Einschätzung dieser Person weggelassen wird. Die entsprechenden Subskalenskalenmediane über alle zehn, acht sowie sieben Expertinnen und Experten sind in Tabelle 5.3 auf Seite 181 einzusehen. Insgesamt fällt auf, dass die Mediane über alle zehn Expertinnen und Experten aller 5 Skalen mit Werten von sechs bis neun von insgesamt elf relativ hoch ausfallen. Vor allem die Bewertung der Items zu den Subskalen Quelle, Rechtfertigung und Entwicklung fällt im Mittel relativ gut aus. Die Bewertung der Items Struktur und Sicherheit fällt weniger gut aus.

Die Itemmediane wurden nun mit den jeweils drei verschiedenen berechneten Subskalenmedianen verglichen. Items wurden nur von den weiteren Analysen entfernt, wenn ihr Skalenmedian in Bezug zu allen drei berechneten Subskalenmedianen geringer ausfiel. Dieses Verfahren wurde gewählt, da sich nach der Expertenbefragung im Rahmen der ersten und zweiten Stichprobe noch weitere Möglichkeiten ergeben, ungeeignete Items zu identifizieren und zu entfernen.

Vor dem Ausschluss eines Items aus dem Inventar, wurden immer die inhaltlichen, itemspezifischen Expertenrückmeldungen berücksichtigt. Denn die Gelegenheit zur inhaltlichen Rückmeldung wurde sehr ausführlich genutzt, sodass die Items auf Basis der Hinweise und Empfehlungen der Experten ausführlich überarbeitet und umformuliert wurden. Die Experten gaben beispielsweise wertvolle Hinweise zu möglichen Formulierungsvarianten und zur Konfundierung einzelner Items mit anderen Skalen. Eine bedeutende Änderung

in der Formulierung, die vorgenommen wurde, bezieht sich auf die Umsetzung der Disziplinspezifität in den einzelnen Items. In allen Items wurde nun über den Begriff *pädagogischer Studientext* – statt wie vorher nur *Studientext* – der Bezug zur relevanten Fachdisziplin hergestellt (Itembeispiel: gw_11 „Beim Lesen eines pädagogischen Studientextes is es das Wichtigste zu erkennen, welche der dargestellten Lösungen die richtige ist.“). Als Resultat der Überarbeitung wurden außerdem sechs Items einer anderen Subskala zugeordnet. Beispielsweise aus Item gw_14 der Subskala *Sicherheit des Wissens* wurde nach der Expertenbefragung Item sw_27 der Subskala *Struktur des Wissens* (sw_27: „Die meisten Autorinnen und Autoren von pädagogischen Studientexten stellen die Themen komplizierter dar, als sie eigentlich sind.“).

TABELLE 5.3: Zusammengefasste Expertenurteile über die Subskalen des 126-Item StEB Testentwurfs

| Subskalen | *MW* | *SD* | *Md* |
|---|---|---|---|
| 10 Experten | | | |
| QW | 7.01 | 3.43 | 8 |
| SW | 6.00 | 3.47 | 6 |
| RW | 6.99 | 3.45 | 8 |
| GW | 6.54 | 3.59 | 7 |
| EW | 7.85 | 3.01 | 9 |
| 8 Experten | | | |
| QW | 6.71 | 3.54 | 8 |
| SW | 5.46 | 3.51 | 6 |
| RW | 6.61 | 3.58 | 7 |
| GW | 5.95 | 3.65 | 6 |
| EW | 7.61 | 3.18 | 9 |
| 7 Experten | | | |
| QW | 6.95 | 3.51 | 8 |
| SW | 5.37 | 3.54 | 5 |
| RW | 6.81 | 3.65 | 8 |
| GW | 5.83 | 3.73 | 6 |
| EW | 7.62 | 3.25 | 9 |

Im Anhang B in den Tabellen B.8 bis B.13 (ab S. 376) ist der genaue Itemwortlaut einzusehen, der aus der Überarbeitung der Items resultierte. Hier geht aus dem Itemcode auch die a priori Zuteilung der Items zu den Subskalen sowie die Formulierungsrichtung der Items hervor. Es wird darüber hinaus aufgezeigt, welche Items auf welchen bestehenden Instrumenten basieren und welche sechs Items einen neuen Itemcode erhielten, da sie einer anderen Subskala zugeordnet wurden.

Wie in Tabelle 5.4 zu sehen, verblieben nach der Überarbeitung der Items und der Itemreduktion 98 Items, von denen 58 Items invertiert gepolt sind.

TABELLE 5.4: Item-Skalen-Zuordnung des 98-Item StEB Inventar Testentwurfs

| Überzeugungs-dimensionen | Skalen-code | Item-anzahl | Itemanzahl invertiert | Itemanzahl abgewandelt |
|---|---|---|---|---|
| Quelle des Wissens | QW | 22 | 17 | 9 |
| Rechtfertigung des Wissens | RW | 26 | 11 | 1 |
| Sicherheit des Wissens | GW | 16 | 12 | 2 |
| Struktur des Wissens | SW | 22 | 14 | 4 |
| Entwicklung des Wissens | EW | 12 | 4 | 4 |
| Gesamt | | 98 | 58 | 20 |

Auf Basis der expertenorientierten Instrumentverbesserung und Itemreduktion wird dargestellt im folgenden Kapitel dargestellt, wie aus diesen verbliebenen 98 Items eine erste Version des StEB Inventars konstruiert, an einer ersten Stichprobe von Lehramtsstudierenden eingesetzt und faktorenanalytisch geprüft wurde.

# 6 Exploratorische Modellbildung (Studie 1)

In einem ersten Schritt wird die EFA anhand einer ersten Stichprobe, dazu eingesetzt, ungeeignete Items und Faktoren zu entfernen und die Faktorstruktur des Testentwurfes zu ermitteln (ebd.). Dieses Kapitel befasst sich dementsprechend mit der *exploratorischen Modellbildung* des StEB Inventars. Hier wird aufgezeigt, wie der durch Expertenurteile optimierte StEB Testentwurf bei einer ersten Stichprobe eingesetzt und unter Einsatz der exploratorischen Faktorenanalyse (EFA) empirisch geprüft wird.

## 6.1 Durchführung

Nach der Expertenbefragung wurde aus den 98 verbliebenen Items eine erste Version des StEB Inventars konstruiert. Welche dieser Items auf spezifischen Items des DEBQ, EBI, SEQ und den Instrumenten von Conley et al. (2004) und Urhahne und Hopf (2004) basieren, ist im Anhang B in den Tabellen B.8 bis B.13 (ab S. 376) einsehbar.

Das Inventar wurde so gestaltet und mit Instruktionen versehen, dass es eine selbständige Bearbeitung durch die Probanden ermöglicht und keiner speziellen Anleitung oder Einführung durch die Versuchsleiterin oder Dozierende bedarf. Dazu beginnt der Fragebogen auf der ersten Seite mit einer kurzen Begrüßung und Bitte um Teilnahme. Auf der zweiten Seite werden demografische Daten erfragt. Das umfasst zunächst die Erstellung eines persönlichen Codes sowie Angaben zu Alter, Geschlecht, Name der Hochschule, angestrebter Abschluss, studierter Schulart, Studiengang, Fächerkombination im Lehramt, Fachsemester, Hochschulsemester und Studienbeginn des aktuellen Studiengangs. Der Code diente dazu festzustellen, ob Probanden wiederholt an der Befragung teilnahmen, um diese gegebenenfalls ausschließen zu können. Auf der dritten Seite des Fragebogens erhalten die Probanden Hinweise zur Beantwortung der sich anschließenden 98 Items. Darüber hinaus wird hier erläutert, was im Rahmen der 98 Statements unter pädagogischen Studientexten zu verstehen ist. Kernstück des StEB Inventars bilden die Seiten 4 bis 9. Mit der Frage: "Wie stark stimmen Sie den Aussagen zu in Bezug auf Ihren Studienbereich Pädagogik/Schulpädagogik/Erziehungswissenschaft?"werden die Probanden dazu angeleitet, zu jedem der 98 als Aussagen formulierten Items eine Einschätzung abzugeben. Dazu wurden die Items mit einer fünfstufigen Rating-Skala mit den Stufen 1 *trifft gar nicht zu*, 2 *trifft eher nicht zu*, 3 *teils teils*, 4 *trifft eher zu* und 5 *trifft voll zu* versehen und über die Skalen hinweg randomisiert. Den Abschluss des Fragebogens bildet ein Anmerkungsfeld. Die Version des

StEB Inventars, die den Probanden der ersten Stichprobe vorgelegt wurde, ist im Anhang A ab Seite 359 einsehbar.

Nach einem Pretest zur Verständlichkeit und Lesbarkeit der Items mit fünf Lehramtsstudierenden der Zielgruppe, aus dem sich jedoch keine Änderungen ergaben, wurde das StEB Inventar im Sommersemester 2016 als Papierversion einer Stichprobe von insgesamt $N = 221$ Lehramtsstudierenden der Universität Passau ($n = 93$) und der Martin-Luther Universität Halle-Wittenberg ($n = 128$) vorgelegt. Die Erhebung erfolgte – nach vorheriger Absprache mit den Dozierenden – in Seminaren und Vorlesungen aus den Studienbereichen der Schulpädagogik, allgemeinen Pädagogik und pädagogischen Psychologie, da die Items sich domänenspezifisch auf Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten beziehen. Die Erhebung in Veranstaltungen des erziehungswissenschaftlichen Bereichs sollte den Studierenden erleichtern, die Instruktion des Fragebogens zu befolgen, die verlangt, die Aussagen der Items ausschließlich auf den Studienbereich Pädagogik, Schulpädagogik und Erziehungswissenschaft zu beziehen. In der Instruktion wurden explizit diese drei Bereiche genannt, um sicherzustellen, dass die befragten Studierenden wissen, auf welche ihrer universitären Veranstaltungen sie die Items beziehen sollen. Das Ausfüllen des Fragebogens nahm ungefähr 15 bis 20 Minuten in Anpruch. Durch die direkte Erhebung mittels einer Papierversion des Fragebogens im Rahmen von Veranstaltungen wurde ein Rücklauf von nahezu 100% erreicht.

## 6.2 Datenbereinigung und deskriptive Datenanalyse

Die Daten zu den 58 invertiert gepolten Items wurden zunächst umkodiert. Die Datensätze von insgesamt vier Versuchspersonen wurden von den weiteren Analysen ausgeschlossen. Zwei der Datensätze wurden wegen zu vieler fehlender Werte ausgeschlossen, da die Studierenden jeweils eine gesamte Seite des Fragebogens nicht bearbeitet hatten. Ein Fragebogen wurde ausgeschlossen, da der Proband viele Items außerhalb der fünfstufigen Skala angekreuzt und zudem viele Aussagen mit Fragezeichen versehen hatte. Der vierte Datensatz wurde wegen einem deutlichen Antwortmuster (Tendenz zur Mitte) ausgeschlossen.

Insgesamt wurden die Daten von $N = 217$ (weiblich 151, männlich 65, k.A. 1) Lehramtsstudierenden der Universität Passau ($n = 91$; 41,9%) und der MLU Halle-Wittenberg ($n = 126$; 58,1%) mit einem durchschnittlichen Alter von 22.6 ($SD = 3.43$) ausgewertet. Die Befragten studieren im Durchschnitt im vierten Fachsemester ($M = 4.31$, $SD = 2.17$) und knapp sechsten Hochschulsemester ($M = 5.95$, $SD = 3.63$). In Bezug auf den angestrebten Studienabschluss studieren 197 der Befragten klassisches Lehramt und 20 Personen eine Kombination aus klassischem Lehramt mit einem Bachelor oder Master. Von den 217 Probanden studiert mit 33.6% (73 Personen) der Großteil Lehramt für Gymnasium, gefolgt von Lehramt für Grundschule und Sekundarschule (siehe Tab. 6.1).

TABELLE 6.1: Häufigkeiten der studierten Schularten ($N = 217$)

| Schultyp | Probanden | % |
|---|---|---|
| Grundschule | 65 | 30,0 |
| Hauptschule | 3 | 1,4 |
| Realschule | 16 | 7,4 |
| Mittelschule | 9 | 4,1 |
| Gymnasium | 73 | 33,6 |
| Förderschule | 24 | 11,1 |
| Sekundarschule | 27 | 12,4 |
| Gesamt | 217 | 100 |

Aus Tabelle 6.2 auf Seite 186 lässt sich entnehmen, dass mit 43,8% der größte Anteil der Befragten angaben, im vierten Fachsemester zu studieren. Den nächst größten Anteil bilden die Sechstfachsemester mit 18,4%. 24 Personen (11,1%) studieren im ersten Fachsemester. Elf der 24 Erstfachsemester befinden sich auch im ersten Hochschulsemester, wie aus Tabelle 6.3 auf Seite 186 hervorgeht. Mit 35,94% befindet sich der größte Anteil der Befragten im vierten Hochschulsemester, gefolgt vom sechsten Hochschulsemester mit 19,82% (siehe Tab. 6.3, S. 186).

TABELLE 6.2: Häufigkeiten und prozentuale Anteile nach Studienbeginn und Fachsemester ($N$ = 217)

| Studienbeginn | Fachsemester | Probanden | % |
|---|---|---|---|
| WiSe 10/11 | 12 | 2 | 0,9 |
| SoSe 11 | 11 | 1 | 0,5 |
| WiSe 11/12 | 10 | 4 | 1,8 |
| SoSe 12 | 9 | 3 | 1,4 |
| WiSe 12/13 | 8 | 8 | 3,7 |
| SoSe 13 | 7 | 4 | 1,8 |
| WiSe 13/14 | 6 | 40 | 18,4 |
| SoSe 14 | 5 | 6 | 2,8 |
| WiSe 14/15 | 4 | 95 | 43,8 |
| SoSe 15 | 3 | 8 | 3,7 |
| WiSe 15/16 | 2 | 22 | 10,1 |
| SoSe 16 | 1 | 24 | 11,1 |
| Gesamt | | 217 | 100 |

TABELLE 6.3: Häufigkeiten und prozentuale Anteile nach Studienbeginn und Hochschulsemester ($N$ = 217)

| Studienbeginn | Hochschulsemester | Probanden | % |
|---|---|---|---|
| WiSe 05/06 | 22 | 1 | 0,5 |
| SoSe 06 | 21 | 1 | 0,5 |
| SoSe 07 | 19 | 1 | 0,5 |
| WiSe 07/08 | 18 | 1 | 0,5 |
| WiSe 08/09 | 16 | 1 | 0,5 |
| SoSe 09 | 15 | 1 | 0,5 |
| WiSe 09/10 | 14 | 5 | 2,3 |
| SoSe 10 | 13 | 1 | 0,5 |
| WiSe 10/11 | 12 | 5 | 2,3 |
| SoSe 11 | 11 | 3 | 1,4 |
| WiSe 11/12 | 10 | 15 | 6,9 |
| SoSe 12 | 9 | 5 | 2,3 |
| WiSe 12/13 | 8 | 14 | 6,5 |
| SoSe 13 | 7 | 5 | 2,3 |
| WiSe 13/14 | 6 | 43 | 19,8 |
| SoSe 14 | 5 | 5 | 2,3 |
| WiSe 14/15 | 4 | 78 | 35,9 |
| SoSe 15 | 3 | 3 | 1,4 |
| WiSe 15/16 | 2 | 15 | 6,9 |
| SoSe 16 | 1 | 11 | 5,1 |
| n.A. | 99 | 3 | 1,4 |
| Gesamt | | 217 | 100 |

Die Tabellen 6.4 und 6.5 zeigen die durch die Lehramtsstudierenden angegebenen Fächer. Deutsch wird als häufigstes Erstfach von 28% der Befragten studiert, gefolgt von Mathematik mit 16% und Englisch mit rund 15%. Auch bezüglich des Zweitfachs sind Mathematik und Deutsch am häufigsten vertreten.

TABELLE 6.4: Häufigkeiten und prozentuale Anteile der studierten Erst- und Zweitfächer ($N$ = 217)

| Erstfach | | | Zweitfach | | |
|---|---|---|---|---|---|
| Fächer | Probanden | % | Fächer | Probanden | % |
| Deutsch | 61 | 28,1 | Mathematik | 49 | 22,8 |
| Mathematik | 35 | 16,1 | Deutsch | 34 | 15,8 |
| Englisch | 32 | 14,7 | Geschichte | 17 | 7,9 |
| Sport | 21 | 9,7 | Geographie | 13 | 6 |
| Biologie | 12 | 5,5 | Religion | 13 | 6 |
| Geographie | 8 | 3,7 | Englisch | 11 | 5,1 |
| Sozialkunde | 9 | 4,1 | Sozialkunde | 10 | 4,7 |
| Religion | 6 | 2,8 | Biologie | 10 | 4,7 |
| Musik | 5 | 2,3 | Wirtschaft | 8 | 3,7 |
| Wirtschaft | 5 | 2,3 | Musik | 7 | 3,3 |
| Sonstige | 28 | 12,9 | Sonstige | 43 | 20 |
| Gesamt | 217 | 100 | | 215 | 100 |

Das am häufigsten gewählte Drittfach ist Mathematik, gefolgt von Sport und Englisch. Als Viertfach wurde am häufigsten Sport gewählt (siehe Tab. 6.5). Da die Befragten Lehramt für unterschiedliche Schularten studieren, unterscheidet sich auch die Anzahl der zu wählenden Fächer. Aus diesem Grund wurde ein Drittfach nur von 89 von 217 Studierenden und ein Viertfach nur von 54 angegeben (siehe Tab. 6.5).

TABELLE 6.5: Häufigkeiten und prozentuale Anteile der studierten Dritt- und Viertfächer ($N$ = 217)

| Drittfach | | | Viertfach | | |
|---|---|---|---|---|---|
| Fächer | Probanden | % | Fächer | Probanden | % |
| Mathematik | 19 | 21,3 | Sport | 17 | 31,5 |
| Sport | 12 | 13,5 | Kunst | 9 | 16,7 |
| Englisch | 10 | 11,2 | Geschichte | 5 | 9,3 |
| Deutsch | 7 | 7,9 | Sachkunde | 5 | 9,3 |
| Musik | 7 | 7,9 | Religion | 4 | 7,4 |
| Sozialkunde | 6 | 6,7 | Mathematik | 3 | 5,6 |
| Geschichte | 5 | 5,6 | Musik | 3 | 5,6 |
| Sachkunde | 5 | 5,6 | Englisch | 2 | 3,7 |
| Sonstige | 18 | 20,2 | Sonstige | 6 | 11,1 |
| Gesamt | 89 | 100 | | 54 | 100 |

### 6.2.1 Deskriptive Analyse der StEB-Items

Die Mittelwerte der StEB-Items liegen zwischen $M$ = 2.23 für Item sw_15_u und $M$ = 4.18 für Items gw_12_u. Die Standardabweichungen der Items reichen von $SD$ = 0.67 (gw_17) bis 1.25 (gw_10_u) (siehe Tab. 6.6, S. 189 und Tab. 6.7, S. 190).

Der Itemmittelwert dient bei ordinalen und intervallskalierten Items als Maß für die psychometrische Itemschwierigkeit (Döring & Bortz, 2016, S. 476 f.). Die Itemschwierigkeit ist insofern interessant, als dass bei einer Instrumententwicklung Items konstruiert werden müssen, mit denen es gelingt, zwischen Personen mit unterschiedlich hohen Merkmalsausprägungen zu differenzieren. Items, denen viele Personen stark zustimmen, verfügen über eine niedrige psychometrische Schwierigkeit. Items, die von vielen Personen eher eine geringe Zustimmung erhalten, sind psychometrisch schwieriger (Döring & Bortz, 2016, S. 476 f.). Zur richtigen Interpretation der Itemschwierigkeit muss die Formulierungsrichtung berücksichtigt werden (Bühner, 2011, S. 235). Die negativ gepolten Items wurden hier bereits umkodiert. Bei umkodierten Items bedeutet ein niedriger Mittelwert also eine hohe Zustimmung, während ein hoher Mittelwert eine geringe Zustimmung zum Item ausdrückt.

Die psychometrisch leichtesten Items sind folglich Item rw_25 („Mir ist es wichtig einzuschätzen, ob das, was ich lese, in sich logisch ist, um beim Lesen eines pädagogischen Studientextes zu Erkenntnissen zu gelangen.“) mit $M$ = 4.13 ($SD$ = .72) und bezüglich der umkodierten Items sw_15_u („In pädagogischen Studientexten werden Dinge oft unnötig verkompliziert.“) mit $M$ = 2.23 $SD$ = 1.04. Diesen Items wurde durchschnittlich am stärksten zugestimmt. Am psychometrisch schwierigsten sind die Items sw_18 („Wirklich anwendbares Wissen muss in pädagogischen Studientexten eher komplex dargestellt werden.“) ($M$ = 2.46, $SD$ = .99), rw_3 („Es ist mir unwichtig, ob mir Inhalte von pädagogischen Studientexten sofort plausibel erscheinen.“) ($M$ = 2.46, $SD$ = .94) und für die umkodierten Items gw_12_u („Wenn in einem pädagogischen Studientext zwei verschiedene Meinungen zu einem Thema dargestellt werden, kann maximal nur eine davon richtig sein.“) ($M$ = 4.18, $SD$ = .75).

Die Itemverteilung ist bei der Bewertung der psychometrischen Schwierigkeit ein wichtiger Aspekt, zumal eine Testkonstruktion auf eine breite Differenzierung des Merkmalsbereichs abzielt und bewusst versucht wird, Items mit einer variierenden Schiefe zu konstruieren. Man geht also gar nicht davon aus, dass eine Normalverteilung der Daten erzielt wird. Das ist auch einer der Gründe, weshalb Bühner (2011, S. 233) einen Test auf Normalverteilung der Items als inhaltlich nicht sinnvoll einschätzt. Darüber hinaus ist das Ergebnis des Tests auf Normalverteilung, bei dem es sich um einen Signifikanztest handelt, stark von der Stichprobengröße abhängig. Bei einer großen Stichprobe würden deshalb schon kleine Abweichungen zu einem signifikanten Ergebnis und einer Ablehnung der Normalverteilung führen Bühner (2011, S. 234).

TABELLE 6.6: Deskriptive Statistik der StEB-Items (Studie 1) – Teil 1

| Items | *N* | *M* | *SD* | *Md* | Min | Max | Range | skew | kurtosis | *SE* |
|---|---|---|---|---|---|---|---|---|---|---|
| qw_5 | 215 | 2.99 | 0.84 | 3 | 1 | 5 | 4 | 0.53 | 0.08 | .06 |
| sw_4 | 215 | 2.89 | 1.09 | 3 | 1 | 5 | 4 | 0.06 | -0.89 | .07 |
| qw_24 | 217 | 3.38 | 0.91 | 3 | 1 | 5 | 4 | -0.31 | -0.22 | .06 |
| sw_22 | 216 | 3.99 | 0.8 | 4 | 1 | 5 | 4 | -0.78 | 1.01 | .05 |
| rw_28 | 217 | 4.11 | 0.91 | 4 | 2 | 5 | 3 | -0.68 | -0.53 | .06 |
| gw_18 | 217 | 3.54 | 0.85 | 4 | 2 | 5 | 3 | -0.14 | -0.61 | .06 |
| rw_27 | 215 | 3.96 | 0.99 | 4 | 1 | 5 | 4 | -0.83 | 0.09 | .07 |
| sw_14 | 215 | 3.93 | 0.84 | 4 | 1 | 5 | 4 | -0.43 | -0.20 | .06 |
| rw_12 | 214 | 3.55 | 0.95 | 4 | 1 | 5 | 4 | -0.58 | 0.02 | .06 |
| qw_11 | 217 | 3.58 | 1.03 | 4 | 1 | 5 | 4 | -0.58 | -0.29 | .07 |
| sw_9 | 216 | 3.69 | 0.89 | 4 | 1 | 5 | 4 | -0.49 | -0.12 | .06 |
| gw_9 | 216 | 3.74 | 0.91 | 4 | 1 | 5 | 4 | -0.51 | -0.19 | .06 |
| ew_4 | 217 | 3.7 | 0.92 | 4 | 2 | 5 | 3 | -0.20 | -0.80 | .06 |
| ew_3 | 216 | 3.59 | 0.85 | 4 | 1 | 5 | 4 | -0.31 | -0.30 | .06 |
| rw_24 | 217 | 3.53 | 0.98 | 4 | 1 | 5 | 4 | -0.38 | -0.39 | .07 |
| sw_10 | 217 | 3.58 | 0.81 | 4 | 1 | 5 | 4 | -0.37 | -0.15 | .06 |
| rw_26 | 216 | 3.38 | 0.89 | 3 | 1 | 5 | 4 | -0.17 | -0.35 | .06 |
| ew_13 | 213 | 4.03 | 0.75 | 4 | 1 | 5 | 4 | -0.52 | 0.48 | .05 |
| rw_13 | 216 | 3.52 | 0.86 | 4 | 1 | 5 | 4 | -0.21 | -0.25 | .06 |
| sw_16 | 217 | 3.88 | 0.93 | 4 | 1 | 5 | 4 | -0.44 | -0.56 | .06 |
| gw_16 | 217 | 3.41 | 0.89 | 3 | 1 | 5 | 4 | -0.26 | -0.17 | .06 |
| sw_18 | 217 | 2.46 | 0.99 | 2 | 1 | 5 | 4 | 0.51 | -0.26 | .07 |
| rw_22 | 216 | 3.69 | 0.82 | 4 | 1 | 5 | 4 | -0.58 | 0.35 | .06 |
| qw_6 | 217 | 3.01 | 0.82 | 3 | 1 | 5 | 4 | 0.22 | 0 | .06 |
| ew_2 | 217 | 3.79 | 0.86 | 4 | 1 | 5 | 4 | -0.72 | 0.62 | .06 |
| sw_24 | 217 | 3.29 | 0.99 | 3 | 1 | 5 | 4 | -0.20 | -0.55 | .07 |
| ew_12 | 217 | 3.82 | 0.9 | 4 | 1 | 5 | 4 | -0.41 | -0.07 | .06 |
| rw_31 | 217 | 3.92 | 0.92 | 4 | 1 | 5 | 4 | -0.76 | 0.32 | .06 |
| rw_23 | 216 | 3.99 | 0.71 | 4 | 2 | 5 | 3 | -0.61 | 0.71 | .05 |
| rw_37 | 217 | 3.97 | 0.81 | 4 | 1 | 5 | 4 | -0.84 | 1.12 | .05 |
| rw_5 | 217 | 3.62 | 0.94 | 4 | 1 | 5 | 4 | -0.70 | 0.24 | .06 |
| ew_15 | 215 | 2.85 | 0.79 | 3 | 1 | 5 | 4 | 0.27 | -0.16 | .05 |
| rw_35 | 217 | 3.07 | 1 | 3 | 1 | 5 | 4 | -0.16 | -0.53 | .07 |
| rw_11 | 216 | 3.97 | 0.83 | 4 | 1 | 5 | 4 | -0.86 | 1.17 | .06 |
| rw_25 | 217 | 4.13 | 0.72 | 4 | 1 | 5 | 4 | -0.88 | 1.74 | .05 |
| rw_3 | 217 | 2.46 | 0.94 | 2 | 1 | 5 | 4 | 0.52 | -0.10 | .06 |
| ew_10 | 217 | 3.51 | 0.87 | 4 | 1 | 5 | 4 | -0.38 | 0.08 | .06 |
| qw_12 | 217 | 3.7 | 0.98 | 4 | 1 | 5 | 4 | -0.68 | -0.03 | .07 |
| ew_8 | 217 | 3.41 | 0.82 | 3 | 1 | 5 | 4 | -0.05 | -0.36 | .06 |
| gw_17 | 217 | 4.08 | 0.67 | 4 | 1 | 5 | 4 | -0.91 | 2.55 | .05 |
| qw_29_u | 217 | 3.61 | 0.87 | 4 | 2 | 5 | 3 | -0.16 | -0.66 | .06 |
| sw_13_u | 217 | 2.78 | 1.03 | 3 | 1 | 5 | 4 | 0.39 | -0.48 | .07 |
| gw_10_u | 217 | 2.88 | 1.25 | 3 | 1 | 5 | 4 | 0.10 | -1.14 | .09 |
| rw_7_u | 216 | 3.4 | 1.23 | 4 | 1 | 5 | 4 | -0.33 | -0.95 | .08 |
| gw_2_u | 215 | 3.31 | 1.18 | 4 | 1 | 5 | 4 | -0.32 | -0.90 | .08 |
| rw_29_u | 217 | 2.98 | 1 | 3 | 1 | 5 | 4 | 0.26 | -0.55 | .07 |
| gw_20_u | 214 | 3 | 1.01 | 3 | 1 | 5 | 4 | 0.10 | -0.62 | .07 |
| sw_1_u | 217 | 2.75 | 1.07 | 3 | 1 | 5 | 4 | -0.03 | -0.59 | .07 |

TABELLE 6.7: Deskriptive Statistik der StEB-Items (Studie 1) – Teil 2

| Items | *N* | *M* | *SD* | *Md* | Min | Max | Range | skew | kurtosis | *SE* |
|---|---|---|---|---|---|---|---|---|---|---|
| rw_36_u | 217 | 3.1 | 1.14 | 3 | 1 | 5 | 4 | -0.09 | -0.61 | .08 |
| qw_26_u | 217 | 3.19 | 1.16 | 3 | 1 | 5 | 4 | -0.16 | -0.81 | .08 |
| rw_33_u | 217 | 3.12 | 1.02 | 3 | 1 | 5 | 4 | 0.17 | -0.7 | .07 |
| sw_21_u | 217 | 3.18 | 0.89 | 3 | 1 | 5 | 4 | -0.03 | -0.63 | .06 |
| qw_13_u | 217 | 3.35 | 0.89 | 3 | 1 | 5 | 4 | -0.17 | -0.59 | .06 |
| qw_21_u | 214 | 2.95 | 0.79 | 3 | 1 | 5 | 4 | 0.55 | 0.18 | .05 |
| qw_15_u | 217 | 2.96 | 0.93 | 3 | 1 | 5 | 4 | -0.1 | -0.19 | .06 |
| sw_11_u | 217 | 2.57 | 0.98 | 3 | 1 | 5 | 4 | 0.02 | -0.7 | .07 |
| rw_19_u | 216 | 3.05 | 0.95 | 3 | 1 | 5 | 4 | 0.07 | -0.65 | .06 |
| gw_4_u | 216 | 2.57 | 0.83 | 3 | 1 | 5 | 4 | 0.34 | 0.02 | .06 |
| qw_20_u | 216 | 2.61 | 0.83 | 2 | 1 | 5 | 4 | 0.79 | 0.58 | .06 |
| qw_25_u | 217 | 3.52 | 0.98 | 4 | 1 | 5 | 4 | -0.22 | -0.51 | .07 |
| sw_7_u | 217 | 3.37 | 0.92 | 4 | 1 | 5 | 4 | -0.45 | -0.26 | .06 |
| qw_22_u | 217 | 3.83 | 0.89 | 4 | 1 | 5 | 4 | -0.49 | -0.1 | .06 |
| gw_13_u | 217 | 3.09 | 1.15 | 3 | 1 | 5 | 4 | -0.15 | -0.89 | .08 |
| qw_19_u | 217 | 2.85 | 1.06 | 3 | 1 | 5 | 4 | 0.27 | -0.59 | .07 |
| rw_2_u | 217 | 2.24 | 0.89 | 2 | 1 | 5 | 4 | 0.53 | -0.03 | .06 |
| gw_15_u | 217 | 4.03 | 0.93 | 4 | 2 | 5 | 3 | -0.51 | -0.82 | .06 |
| sw_17_u | 217 | 2.54 | 0.94 | 2 | 1 | 5 | 4 | 0.49 | -0.03 | .06 |
| gw_1_u | 216 | 3.39 | 0.9 | 3 | 1 | 5 | 4 | -0.06 | -0.36 | .06 |
| qw_14_u | 216 | 3.06 | 0.97 | 3 | 1 | 5 | 4 | -0.19 | -0.38 | .07 |
| sw_23_u | 217 | 2.48 | 0.76 | 2 | 1 | 5 | 4 | 0.36 | 0 | .05 |
| qw_8_u | 217 | 3.37 | 1.18 | 4 | 1 | 5 | 4 | -0.36 | -0.79 | .08 |
| sw_2_u | 216 | 2.45 | 0.96 | 2 | 1 | 5 | 4 | 0.37 | -0.49 | .07 |
| gw_3_u | 216 | 2.74 | 1.05 | 3 | 1 | 5 | 4 | 0.35 | -0.44 | .07 |
| rw_30_u | 217 | 3.97 | 0.91 | 4 | 1 | 5 | 4 | -0.66 | -0.17 | .06 |
| qw_23_u | 217 | 2.71 | 1.08 | 3 | 1 | 5 | 4 | 0.2 | -0.75 | .07 |
| rw_18_u | 216 | 3.36 | 0.99 | 3 | 1 | 5 | 4 | -0.05 | -0.67 | .07 |
| sw_15_u | 216 | 2.23 | 1.04 | 2 | 1 | 5 | 4 | 0.48 | -0.53 | .07 |
| gw_19_u | 216 | 3.55 | 0.82 | 4 | 1 | 5 | 4 | -0.06 | -0.3 | .06 |
| sw_8_u | 217 | 3.66 | 0.81 | 4 | 1 | 5 | 4 | -0.54 | 0.37 | .05 |
| ew_11_u | 216 | 4.02 | 0.86 | 4 | 1 | 5 | 4 | -0.9 | 0.82 | .06 |
| qw_30_u | 216 | 3.29 | 0.89 | 3 | 1 | 5 | 4 | 0.11 | -0.64 | .06 |
| sw_19_u | 217 | 3.83 | 0.93 | 4 | 1 | 5 | 4 | -0.6 | 0.03 | .06 |
| rw_4_u | 214 | 3.29 | 0.98 | 3 | 1 | 5 | 4 | 0.18 | -0.79 | .07 |
| gw_11_u | 217 | 3.13 | 1.03 | 3 | 1 | 5 | 4 | 0.1 | -0.8 | .07 |
| ew_9_u | 217 | 3.48 | 0.88 | 4 | 1 | 5 | 4 | -0.17 | -0.38 | .06 |
| qw_10_u | 217 | 3.21 | 0.84 | 3 | 1 | 5 | 4 | 0.24 | -0.37 | .06 |
| ew_5_u | 217 | 3.1 | 0.83 | 3 | 1 | 5 | 4 | 0.16 | -0.18 | .06 |
| rw_32_u | 217 | 2.5 | 1.02 | 3 | 1 | 5 | 4 | 0.07 | -0.73 | .07 |
| qw_31_u | 217 | 3.15 | 1.11 | 3 | 1 | 5 | 4 | -0.03 | -0.82 | .08 |
| gw_12_u | 217 | 4.18 | 0.75 | 4 | 2 | 5 | 3 | -0.57 | -0.20 | .05 |
| rw_8_u | 217 | 2.88 | 0.91 | 3 | 1 | 5 | 4 | 0.26 | -0.08 | .06 |
| gw_7_u | 215 | 2.27 | 0.93 | 2 | 1 | 5 | 4 | 0.72 | 0.39 | .06 |
| ew_16_u | 216 | 3.47 | 0.76 | 3 | 1 | 5 | 4 | 0.4 | 0.03 | .05 |
| sw_20_u | 216 | 3.5 | 0.78 | 3 | 1 | 5 | 4 | 0.07 | -0.12 | .05 |
| qw_28_u | 217 | 4.06 | 0.79 | 4 | 2 | 5 | 3 | -0.44 | -0.48 | .05 |
| sw_3_u | 216 | 3.9 | 0.83 | 4 | 1 | 5 | 4 | -0.44 | -0.1 | .06 |
| qw_27_u | 216 | 3.25 | 0.80 | 3 | 2 | 5 | 3 | 0.39 | -0.2 | .05 |
| sw_27_u | 216 | 2.57 | 1.09 | 3 | 1 | 5 | 4 | 0.22 | -0.66 | .07 |

Anhand der Minimal- und Maximalwerte in den Tabellen 6.6 und 6.7 (S. 189, 190) ist gut zu erkennen, dass bei neun Überzeugungsitems nicht die volle Breite des Antwortspektrums genutzt wird. Die betroffenen Items sind:

**qw_29_u:** Wenn mir beim Lesen eines pädagogischen Studientextes auffällt, dass dieser meiner Erfahrung nicht entspricht, dann bin ich mir sicher, mich wohl geirrt zu haben.

**rw_28:** Ich bilde mir meine Meinung zu pädagogischen Themen nicht auf Basis eines einzelnen Studientextes.

**gw_18:** Wenn in einem pädagogischen Studientext mehrere unterschiedliche Ansichten beschrieben werden, dann kann je nach Kontext variieren, welche Sichtweise gerade die bessere ist.

**ew_4:** Es ist wahrscheinlich, dass der Großteil der Erkenntnisse in den pädagogischen Studientexten in zehn Jahren ganz anders aufgefasst wird.

**gw_15_u:** Was in einem pädagogischen Studientext steht, trifft in jedem Fall zu.

**rw_23:** Um einschätzen zu können, ob die in einem pädagogischen Studientext dargestellte Meinung verlässlich ist, ist es wichtig, dass ich versuche Verbindungen zu meinem Vorwissen herzustellen.

**gw_12_u:** Wenn in einem pädagogischen Studientext zwei verschiedene Meinungen zu einem Thema dargestellt werden, kann maximal nur eine davon richtig sein.

**qw_28_u:** Man kann in der Regel alles glauben, was man in pädagogischen Studientexten liest.

**qw_28_u:** Wenn ein pädagogischer Studientext ein bestimmtes Vorgehen empfiehlt, finde ich es wichtig, mich daran zu halten.

Die ungenutzte Antwortkategorie ist ein Indikator einer zu hohen oder niedrigen psychometrischen Itemschwierigkeit, oder aber ein Hinweis darauf, „dass das Antwortformat für die Zielpopulation zu differenziert ist“ (Bühner, 2011, S. 233). Die oben beschriebenen neun Items werden deshalb von der faktorenanalytischen Auswertung ausgeschlossen.

### 6.2.2 Analyse fehlender Werte

Da die exploratorische Faktorenanalyse keine fehlenden Werte erlaubt (Baglin, 2014), ist der angemessene Umgang mit fehlenden Werten eine zentrale Voraussetzung für die weiteren Analysen. Nach Rubin (1976) werden fehlende Daten nach dem zugrundeliegenden Mechanismus des Fehlens klassifiziert: *missing completely at random* (MCAR), *missing at random* (MAR) und *missing not at random* (MNAR). Die MCAR Annahme lässt sich mittels des Little Tests prüfen, während die MAR Annahme sich nicht direkt testen lässt (Enders, 2010, S. 19 f.).

Die 98 Items zu den epistemologischen Überzeugungen wurden zunächst mittels des Little-Tests (Little & Rubin, 1987) auf eine zugrundeliegende Systematik der fehlenden Werte (*missing-completely-at-random*, MCAR) überprüft (R Paket: BaylorEdPsych). Der Little-Test prüft, ob die fehlenden Werte im Datensatz systematisch verteilt sind, das heißt ob der Mechanismus, der hinter dem Fehlen von Daten steht, mit den fehlenden Werten selbst oder aber mit den beobachteten Werten in Zusammenhang steht. Mit $\chi^2 = 2672.049, p = .134$ konnte gezeigt werden, dass keine Verletzung der MCAR-Bedingung vorliegt (Little & Rubin, 2002) und den fehlenden Werten somit keine Systematik zugrunde liegt. Von den 217 Datensätzen sind 187 bezüglich der Überzeugungsitems vollständig. Der prozentuale Anteil fehlender Werte je Überzeugungsitem ist mit maximal 1,8 Prozent bei Item ew_13 sehr gering. Insgesamt haben vier Personen Item ew_13 nicht beantwortet.

Es gibt drei übergeordnete Arten des Umgangs mit fehlenden Werten: Ausschluss (*deletion*), einfache Imputation (*single imputation*) und multiple Imputation (Enders, 2010). In Bezug auf *deletion* unterscheidet man zwischen dem entweder listenweisen oder paarweisen Ausschluss der Fälle mit fehlenden Werten (Enders, 2010, S. 39 f.). Bei den Methoden der einfachen Imputation wird der fehlende Wert in den Daten jeweils mit einem einzelnen Wert ersetzt (Enders, 2010, S. 42f.). Unter diesen Ansatz fallen die folgenden Methoden: *Arithmetic mean imputation*, *Regression imputation*, *Stochastic regression imputation*, *Hot-deck imputation*, *Similar response pattern imputation*, *Averaging the available items* und *Last observation carried forward*. Die multiple Imputation generiert dagegen verschiedene Kopien des Datensatzes und ersetzt in jedem der Datensätze die fehlenden Werte mit unterschiedlichen, plausiblen Schätzwerten.

Aufgrund der geringen Anzahl fehlender Werte und des gleichzeitigen Vorliegens von MCAR wurde für den Umgang mit fehlenden Werten bei den 98 Items zu den epistemologischen Überzeugungen die *arithmetic mean imputation* als Methode zum Umgang mit den fehlenden Werten gewählt. Die fehlenden Itemwerte wurden jeweils durch den univariaten (unbedingten), gerundeten (ohne Dezimalstelle) Stichprobenmittelwert der betroffenen Variable ersetzt. Dieses Imputationsverfahren wurde der gebräuchlichen ipsativen Mittelwertsimputation vorgezogen, bei der ein Mittelwert über die übrigen Items der jeweiligen Person gebildet und für den fehlenden Itemwert eingesetzt wird, da dieses Verfahren nur für eindimensionale Items angemessen ist (Schafer & Graham, 2002, S. 157). Im vorliegenden Fall wird jedoch explizit eine mehrdimensionale Itemsstruktur angenommen, da die Überzeugungsitems mit dem Ziel entwickelt wurden, fünf verschiedenen Skalen anzugehören und dementsprechend unterschiedliche Aspekte epistemologischer Überzeugungen zu erfassen.

## 6.3 Faktorenanalytische Auswertung der ersten Stichprobe

Wie in der Instrumententwicklung nach der klassischen Testtheorie (KTT) üblich, wurden die Daten der ersten Stichprobe mittels der exploratorischen Faktorenanalyse ausgewertet, um die Faktorstruktur des zukünftigen Fragebogens zu untersuchen und die Itemauswahl zu optimieren (Brown, 2015, S. 10). Die EFA wurde in R Statistics, Version 3.4.0 (Team, 2017) umgesetzt.

### 6.3.1 Prüfung der Voraussetzung für die EFA

Generell gilt in Bezug auf die Faktorenanalyse, dass größere Stichproben zu stabileren Faktorlösungen führen (Bühner, 2011, S. 345). Nach den Vorgaben von Mundform et al. (2005) zu den Abhängigkeiten zwischen Stichprobengröße, Itemkommunalitäten und Itemanzahl je Faktor gilt, dass die Stichprobe für die EFA umso kleiner sein kann (bis minimal $N = 55$), je höher die Itemanzahl je Faktor und die Itemkommunalitäten sind (Bühner, 2011, S. 345). Die vorliegende Stichprobengröße von $N = 217$ ist, selbst wenn man zunächst einmal von niedrigen Kommunalitäten und einer sich noch reduzierenden Itemanzahl je Faktor ausgeht, ausreichend (vgl. Kapitel 4.3.1, Tab. 4.1, S. 156).

Grundlage der exploratorischen Faktorenanalyse stellt eine Korrelationsmatrix dar. Die Interitemkorrelationen reichen von -.38 bis .68. Die Korrelationsmatrix wurde zunächst inspiziert, um die Muster der Itembeziehungen aufzudecken, denn die Anwendung der Faktorenanalyse ist nur dann angemessen, wenn die Items ausreichend hoch korrelieren (Bühner, 2011, S. 343). Variablen, die mit keiner anderen Variable korrelieren, sollten von den Analysen ausgeschlossen werden (Field et al., 2012, S. 774). Jedes Item korreliert signifikant mit mindestens einem anderen Item. Es gibt jedoch Items, die keine Zusammenhänge mit anderen Items von mindestens $r =| .30 |$ aufweisen. Dies betrifft die sieben Items: sw_4, rw_22, rw_35, rw_3, rw_7_u, qw_22_u, sw_23_u. Diese sieben Items müssen im Verlauf der Faktorenanalysen im Zusammenhang mit der Höhe der Kommunalitäten und Itemladungen noch genauer analysiert werden, bevor eine Entscheidung über ihren etwaigen Ausschluss erfolgen kann. Zu hohe Korrelationskoeffizienten von ab $r =| .90 |$ (Field et al., 2012, S. 774) liegen nicht vor.

Ausreißerwerte können einen stark mindernden oder erhöhenden Einfluss auf die Zusammenhangsmaße haben und sollten deshalb dringend berücksichtigt werden (Bühner, 2011, S. 343). Zudem weist Bühner (2011, S. 343) darauf hin, dass inhaltliche Gründe und nicht rein statistische den Ausschlag geben sollten, Personen aus dem Datensatz zu entfernen. In dem vorliegenden Fall wurden im Rahmen der Datenaufbereitung (siehe Kap. 6.2) vier Fälle aus dem Datensatz entfernt. Gleichzeitig ist durch das genutzte fünfstufige Antwortformat der Items sichergestellt, dass sich Ausreißerwerte nicht sonderlich stark auswirken (Bühner, 2011, S. 343).

Der Bartlett-Test und der Kaiser-Meyer-Olkin-Koeffizient (KMO) stellen wichtige Kriterien für die Durchführbarkeit einer Faktorenanalyse dar und beziehen sich ebenfalls auf die Prüfung der Korrelationsmatrix. Die Measure-of-Sampling-Adequacy-Koeffizienten (MSA) geben dagegen Auskunft über die Eignung der einzelnen Items. Sie werden vergleichbar mit

dem KMO-Koeffizienten berechnet und bewertet (Bühner, 2011, S. 347). Die Bewertung des KMO- und der MSA-Maße erfolgt nach der in Kapitel 4.3.1 dargestellten Interpretationshilfe in Tabelle 4.2 (S. 157) nach Kaiser und Rice (1974). Das KMO-Maß bestätigt mit KMO = .72 eine mittlere Eignung (siehe Tabelle 4.2), der gesamten Korrelationsmatrix zur Durchführung einer Faktorenanalyse, d.h. kommt zu dem Schluss, dass substantielle Korrelationen vorliegen. Die individuellen, itemspezifischen MSA-Indizes bewegen sich zwischen .44 und .86. Die MSA-Werte von drei Items (sw_18[1] rw_3[2] und sw_23_u[3]) sind mit MSA < .50 als inakzeptabel und für die faktorenanalytische Auswertung ungeeignet einzustufen. Die Items rw_3 und sw_23_u sind oben schon dadurch aufgefallen, dass sie keine Zusammenhänge mit anderen Items von mindestens $r =| .30 |$ aufweisen.

Im Sinne von Bühner (2011, S. 348) wurde geprüft, ob die entsprechenden Items im Vergleich mit den übrigen Items eine ungewöhnliche Verteilung aufweisen, da dies die geringe Korrelation mit anderen Items, die durch das MSA-Maß geprüft werden, erklären würde. Die Prüfung der Histogramme der Überzeugungsitems zeigt jedoch, dass die betroffenen drei Items bei weitem nicht die einzigen mit einer schiefen Verteilung sind. Die drei Items werden deshalb aufgrund ihrer niedrigen MSA-Werte aus den weiteren Analysen ausgeschlossen. Das nun erneut berechnete KMO-Maß verbessert sich durch den Ausschluss der drei Items leicht und beträgt nun KMO = .73. Auch die itemspezifischen MSA-Maße nehmen zu.

Der Bartlett-Test auf Sphärizität ist als Signifikanztest von der Stichprobengröße abhängig und stellt somit bei signifikantem Ergebnis nur eine Minimalbedingung für die Durchführung einer Faktorenanalyse dar. Sofern multivariate Normalverteilung gegeben ist, kann mittels dieses Tests geprüft werden, ob die Korrelationsmatrix faktorisiert werden kann. Dazu wird getestet, ob die Korrelationen signifikant von null abweichen (Bühner, 2011, S. 347). Das Vorliegen multivariater Normalverteilung wurde mithilfe des Mardia-Tests überprüft (Mardia, 1970). Er ergab sowohl für Schiefe als auch für Kurtosis der Items signifikante Ergebnisse ($p < .05$). Aus diesem Grund kann multivariate Normalverteilung für die Überzeugungsitems nicht angenommen werden und somit ist eine Berechnung des Bartlett-Tests nicht zulässig.

### 6.3.2 Vorbereitung der EFA

Bezüglich des Bestimmung der geeigneten Faktorenanzahl empfiehlt Bühner (2011, S. 349), zunächst die Anzahl von Faktoren zu extrahieren, die durch das hypothetische Modell vorgegeben wird. Das spricht im vorliegenden Fall für die Extraktion von fünf Faktoren, da dies der Anzahl an angenommenen Skalen entspricht. Neben den inhaltlichen Überlegungen gibt es, wie bereits in Kapitel 4.3.1 (ab S. 153) beschrieben, deskriptivstatistische

[1] sw_18: „Wirklich anwendbares Wissen muss in pädagogischen Studientexten eher komplex dargestellt werden.“, MSA = .44

[2] „Es ist mir unwichtig, ob mir Inhalte von pädagogischen Studientexten sofort plausibel erscheinen.“, MSA = .46

[3] sw_23_u: „Es ist wichtig beim Lesen eines pädagogischen Studientextes nach den spezifischen Fakten zu suchen.“, MSA = .47

und inferenzstatistische Methoden zur Bestimmung der Faktorenanzahl (Bühner, 2011). Von den deskriptivstatistischen Methoden eignen sich die Weiterentwicklungen des Scree-Tests, das Optimal Coordinate- (OC) und das Acceleration Factor-Verfahren (AF) besser, als das Kaiser-Kriterium und der Scree-Test (Ruscio & Roche, 2012). Zu den empfohlenen inferenzstatistischen Methoden zur Bestimmung der adäquaten Faktorenanzahl (vgl. Kapitel 4.3.1, S. 153) gehört die Parallelanalyse (PA, Horn-Kriterium,Horn, 1965) und die MAP Methode (Minimum Average Partial, Velicer, 1976). Die empfohlene Faktorenanzahl der OC- und AF-Methoden lassen sich mittels des speziell dafür entwickelten R Pakets *nFactors* (Raiche & Magis, 2010) ermitteln.

Abbildung 6.1 kombiniert in ihrer Darstellung den Scree-Plot (4 Faktoren), das Kaiser-Kriterium (24 Faktoren), die Parallelanalyse (7 Faktoren), den Acceleration Faktor (4 Faktoren) und Optimal Coordinates (7 Faktoren). Die MAP-Methode ermittelt eine Anzahl von 6 Faktoren (nicht in Abbildung dargestellt).

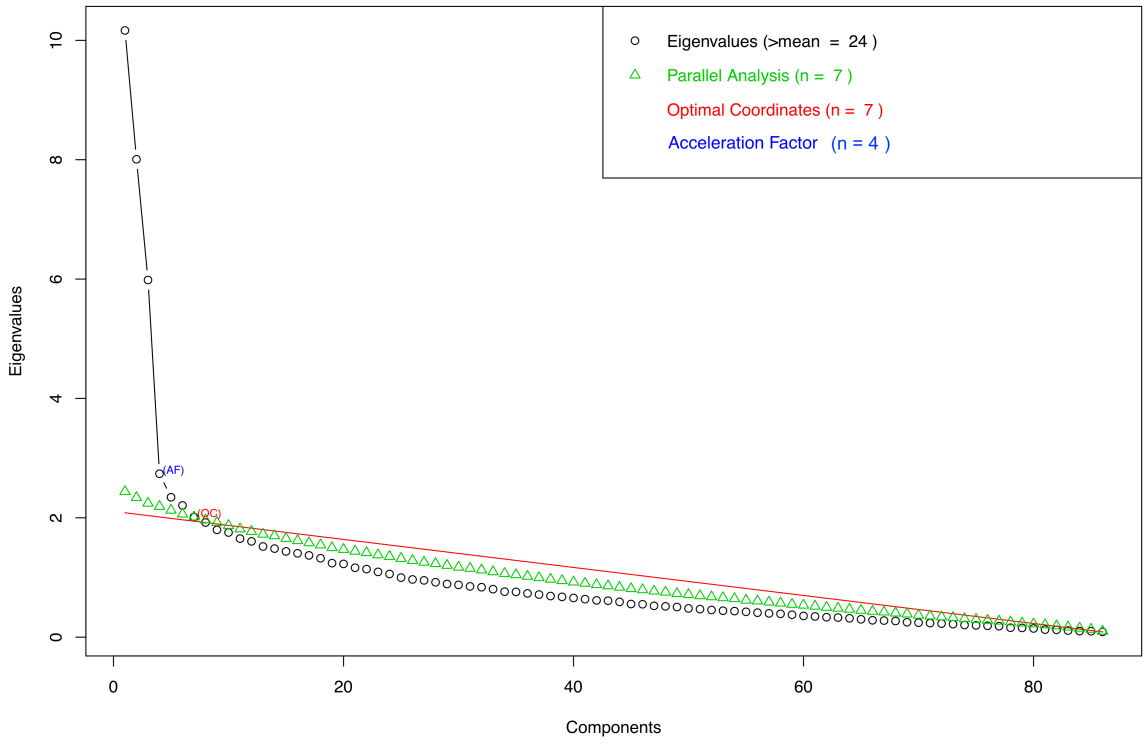


ABBILDUNG 6.1: Kombinierte Darstellung verschiedener Tests zur Bestimmung der geeigneten Faktorenanzahl

Wie erkennbar ist, führen die verschiedenen Verfahren zu unterschiedlichen Ergebnissen bezüglich der Faktorenanzahl. Besondere Berücksichtigung sollten jedoch die Ergebnisse der Parallelanalyse und der MAP Methode finden, da diese in Simulationsstudien zuverlässig die korrekte Anzahl an zugrundeliegenden Dimensionen identifizieren konnten (Hubbard & Allen, 1987; Ruscio & Roche, 2012; Velicer et al., 2000; Zwick & Velicer, 1986) und deshalb empfohlen werden (Lance et al., 2006; Velicer et al., 2000). Im vorliegenden Fall identifiziert die Parallelanalyse (PA) eine Faktorenanzahl von 7, während die

MAP-Methode 6 Faktoren vorschlägt. Beide Ergebnisse entsprechen wiederum nicht der inhaltlich-theoretischen Überlegung, die eine Extraktion von 5 Faktoren nahe legt. Hervorzuheben ist, dass für die Faktorenanzahl maßgeblich sein muss, ob sich die resultierenden Faktoren inhaltlich sinnvoll interpretieren lassen (Bühner, 2011). Um dies zu prüfen, muss jeweils eine EFA für die Faktorenanzahl von 4 (AF), 5 (inhaltlich-theoretisch angenommen), 6 (MAP) und 7 (PA) Faktoren gerechnet, verglichen und auf inhaltliche Interpretierbarkeit hin geprüft werden.

Wie in Kapitel 4.3.1 (ab S. 153) bereits beschrieben, gibt es viele verschiedene Extraktionsmethoden. Zu den bewährtesten Methoden zählt die Maximum-Likelihood-Faktorenanalyse. Allerdings sollte die Auswahl der Extraktionsmethode immer auf Basis der zugrundeliegenden Verteilung der Daten getroffen werden (Baglin, 2014, S. 3). Die Maximum-Likelihood-Methode setzt multivariate Normalverteilung der Daten voraus (Baglin, 2014, S. 3). Wie oben im Zusammenhang mit dem signifikanten Mardia-Test ($p < .05$) berichtet wurde, liegt für die Überzeugungsitems keine multivariate Normalverteilung vor. Als angemessene Extraktionsmethode wurde somit die Hauptachsenanalyse (Principal Axis Factoring, PAF) gewählt, da dieser keine Verteilungsannahmen zugrunde liegen (Fabrigar et al., 1999).

Ist eine Skala mehrdimensional, wie im vorliegenden Fall angenommen wird, muss zur Verbesserung der Interpretierbarkeit des resultierenden Faktormodells eine Faktorrotation (siehe Kapitel 4.3.1, ab S. 153) vorgenommen werden (Beavers et al., 2013). Ist von Zusammenhängen zwischen den Faktoren auszugehen, wie in den Sozialwissenschaften meist der Fall (Basto & Pereira, 2012, S. 3), führt ein obliques Rotationsverfahren zu genaueren Faktorlösungen (Baglin, 2014). Die zwei wichtigsten obliquen Verfahren sind Promax und Direct-Oblimin (Bühner, 2011, S. 338). Da gezeigt werden konnte (Costello & Osborne, 2005; Fabrigar et al., 1999), dass die oblique Rotation zu fast identischen Ergebnissen führt wie eine orthogonale Rotation, wenn die Faktoren wider Erwarten doch unkorreliert sind, sind oblique Verfahren insgesamt zu bevorzugen (Basto & Pereira, 2012, S. 3). Im vorliegenden Fall wurde die Promax-Rotation gewählt (vgl. hierzu Bühner, 2011, S. 338).

### 6.3.3 Ergebnisse der EFA

Es wurden insgesamt fünf Hauptachsenanalysen jeweils mit obliquer Promax-Rotation unter Extraktion von 3, 4, 5, 6 und 7 Faktoren gerechnet. Die Extraktion von 3 Faktoren wurde zusätzlich vorgenommen, um eine Überschätzung der Faktorenanzahl auszuschließen. Der Vergleich der Faktorlösungen ergab, dass allein die Extraktion von 5 Faktoren zu einem inhaltlich-theoretisch interpretierbaren Ergebnis führt. Aus diesem Grund und weil eine Darstellung aller Ergebnisse dieser Faktorenanalysen den Rahmen dieser Arbeit sprengen würde, beschränken sich die folgenden Beschreibungen auf die inhaltlich bedeutsamen Ergebnisse der Hauptachsenanalyse unter Extraktion von 5 Faktoren.

Im folgenden werden die Ergebnisse der Residualanalyse sowie die Kommunalitäten und Ladungen vorgestellt. Wie aus dem Histogramm der Residuen (Abbildung 6.2) zu entnehmen ist, scheinen die Residuen ungefähr normalverteilt zu sein und nicht über Ausreißer zu verfügen. Positiv ist auch, dass mit 0,43% weniger als 50% der Residuen über absolute Werte > .05 verfügen (Field et al., 2012, S. 787). Der RMSR (*root mean square of the residuals*) = .05, bzw. korrigiert für die Freiheitsgrade .06).

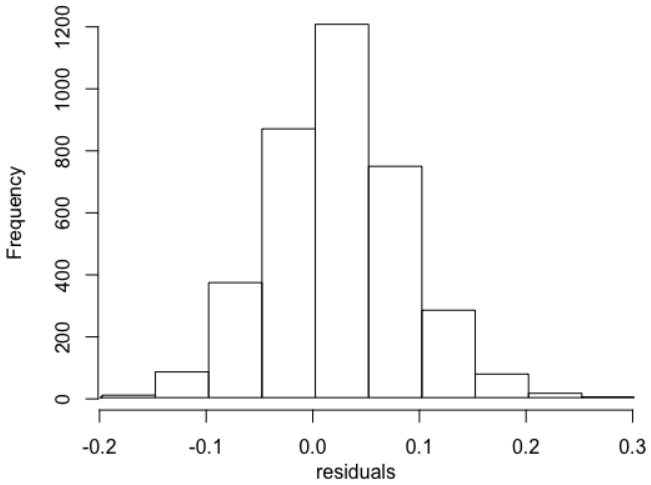


ABBILDUNG 6.2: Histogramm der Residuen der 5-Faktorlösung

Die Tabelle 6.8 und deren Fortführung Tabelle 6.9 auf den Seiten 198 und 199 stellen die Mustermatrix mit den Faktorladungen und Itemkommunalitäten der 5-Faktorlösung dar. Sehr geringe Ladungen von $\lambda < .30$ sind der Übersichtlichkeit halber nicht dargestellt. Die Kommunalität $h^2$ gibt an, wie viel Varianz eines Items durch alle extrahierten Faktoren erklärt wird und stellt eine Mindestschätzung der Itemreliabilitäten dar (Bühner, 2011, S. 344). Die Höhe der Kommunalitäten ist normalerweise abhängig von der Anzahl der zu extrahierenden Faktoren, wobei mehr Faktoren zu höheren Kommunalitäten führen (Bühner, 2011, S. 358). Wie aus Tabellen 6.8 und 6.9 (S. 198 und 199) zu entnehmen ist, verfügen 15 der 86 Items über mittlere Kommunalitäten ($h^2$ = .40 - .60), während bei dem Großteil niedrige Kommunalitäten ($h^2$ = .20 - .39) vorherrschen (Bühner, 2011, S. 345). 14 Items (siehe Tab. 6.8, 6.9, S. 198 f.) verfügen über sehr niedrige Kommunalitäten < .20.

TABELLE 6.8: Mustermatrix mit Faktorladungen $\lambda$ und zugehörigen Kommunalitäten der 5-Faktorlösung (89-Item StEB Inventar Version) – Teil 1

| Item | Faktor 1 | Faktor 2 | Faktor 3 | Faktor 4 | Faktor 5 | $h^2$ |
|---|---|---|---|---|---|---|
| rw_33_u | .80 | | | | | .51 |
| qw_20_u | .76 | | | | | .43 |
| qw_25_u | .70 | | | | | .54 |
| qw_19_u | .62 | | | | | .39 |
| qw_21_u | .60 | | | | | .37 |
| gw_1_u | .56 | | | | | .39 |
| qw_23_u | .55 | | | | | .31 |
| qw_26_u | .54 | | | | | .31 |
| qw_13_u | .53 | | | | | .43 |
| rw_8_u | .51 | | | | | .32 |
| gw_4_u | .50 | | | | | .31 |
| rw_18_u | .48 | | | | | .36 |
| rw_29_u | .47 | | | | | .34 |
| qw_10_u | .47 | | | | | .25 |
| gw_7_u | .46 | | | | | .31 |
| qw_5 | .44 | -.32 | | | | .29 |
| sw_21_u | .41 | | | | | .24 |
| qw_8_u | .40 | | .37 | | | .28 |
| rw_30_u | .38 | | .35 | | | .26 |
| sw_13_u | .35 | | | | | .26 |
| qw_6 | .34 | | | | | .20 |
| qw_31_u | .34 | | | | | .23 |
| sw_15_u | | .67 | | | | .52 |
| sw_1_u | | .66 | | | | .44 |
| sw_27_u | | .61 | | | | .42 |
| gw_13_u | | .57 | | | | .38 |
| gw_3_u | | .57 | | | | .33 |
| sw_2_u | | .57 | | | | .45 |
| gw_2_u | | .56 | | | | .34 |
| sw_3_u | | .54 | | | | .32 |
| sw_11_u | | .53 | | | | .32 |
| sw_17_u | .34 | .53 | | | | .46 |
| qw_15_u | -.33 | .51 | | | | .44 |
| rw_32_u | | .49 | | | -.48 | .50 |
| rw_36_u | | .44 | | | | .32 |
| ew_11_u | | .43 | | | | .30 |
| qw_14_u | -.32 | .42 | | | -.42 | .46 |
| sw_8_u | | .41 | | | | .17 |
| gw_11_u | .32 | .39 | | | | .38 |
| sw_20_u | | .38 | | | | .24 |
| gw_10_u | | .35 | | | | .20 |
| sw_7_u | | .32 | | | | .15 |
| qw_11 | | .32 | .62 | | | .55 |

TABELLE 6.9: Mustermatrix mit Faktorladungen $\lambda$ und zugehörigen Kommunalitäten der 5-Faktorlösung (89-Item StEB Inventar Version) – Teil 2

| Item | Faktor 1 | Faktor 2 | Faktor 3 | Faktor 4 | Faktor 5 | $h^2$ |
|---|---|---|---|---|---|---|
| gw_9 | | | .58 | | | .35 |
| sw_9 | | | .54 | | | .33 |
| rw_27 | | | .54 | | | .36 |
| rw_5 | .32 | | .51 | | | .43 |
| gw_16 | | | .51 | | | .25 |
| rw_12 | | | .49 | | | .37 |
| qw_12 | | | .48 | | | .26 |
| rw_26 | | | .48 | | | .31 |
| rw_11 | | | .47 | | | .32 |
| rw_24 | | | .44 | | | .25 |
| sw_22 | | | .44 | | | .25 |
| sw_10 | | | .42 | | | .19 |
| gw_17 | | | .40 | | | .31 |
| ew_10 | | | .40 | | | .23 |
| rw_22 | | | .36 | | | .22 |
| qw_24 | | | .33 | | | .15 |
| rw_7_u | | | -.32 | | | .13 |
| ew_13 | | | | .61 | | .40 |
| ew_12 | | | | .59 | | .37 |
| ew_2 | | | | .44 | | .25 |
| ew_16_u | | | | .42 | | .28 |
| ew_5_u | | | | .39 | -.32 | .33 |
| ew_3 | | | | .37 | | .32 |
| sw_19_u | | | | .37 | | .25 |
| ew_8 | | | | .35 | | .21 |
| ew_9_u | | | | .33 | | .23 |
| sw_14 | | | | .31 | | .17 |
| rw_2_u | | | | | -.48 | .27 |
| rw_37 | | | .37 | | .39 | .36 |
| rw_25 | | | | | .38 | .28 |
| rw_13 | | | .32 | | .33 | .34 |
| rw_4_u | | | | | .31 | .26 |
| gw_19_u | .30 | | | | | .31 |
| gw_20_u | | | | | | .22 |
| rw_31 | | | | | | .13 |
| rw_19_u | | .30 | | | | .19 |
| ew_15 | | | | | | .16 |
| qw_30_u | | | | | | .09 |
| sw_24 | | | | | | .09 |
| rw_35 | | | | | | .12 |
| sw_16 | | | | | | .25 |
| qw_22_u | | | | | | .13 |
| sw_4 | | | | | | .13 |

Davon werden die Items sw_24 und qw_30_u, mit ihren Kommunalitäten von $h^2 < .10$, von den fünf Faktoren besonders unzureichend erfasst (siehe Tab. 6.9, S. 199). Das bedeutet, dass der Anteil aufgeklärter Varianz für diese beiden Items kleiner als 10% ist (Bühner, 2011, S. 358). Inwiefern die Höhe der Kommunalität die Reduktion der Items beeinflusst, wird im Zusammenhang mit den Faktorladungen diskutiert. Tabachnick und Fidell (2001, S. 358) geben an, dass Ladungen über einem Minimum von $\lambda =| .32 |$ liegen sollten, um als bedeutsam angesehen zu werden. Ladungen über $| .60 |$ können als hoch bezeichnet werden (P. Kline, 1994). Somit liegen für vier der fünf Faktoren, wie aus den Tabellen 6.8 und 6.9 (S. 198 und 199) ersichtlich, hohe Faktorladungen vor. Elf Items laden dagegen auf keinem der Faktoren mit mindestens .32. Es liegen einige Nebenladungen vor. Das bedeutet, dass einige Items zusätzlich zu ihrer Hauptladung auf einem oder mehreren weiteren Faktoren substanziell laden ($\lambda \geq| .32 |$; Costello & Osborne, 2005).

### 6.3.4 Itemreduktion und inhaltliche Interpretation der Faktoren

Bei der Auswahl beizubehaltender Items werden die Höhe der Kommunalitäten und Ladungen sowie der Inhalt der Items berücksichtigt. Zunächst werden die eindeutig ungeeigneten Items aussortiert. Das betrifft solche, die keine bedeutsamen Ladungen aufweisen sowie über geringen Kommunalitäten von $h^2 < .20$ verfügen, d.h. von den fünf Faktoren besonders unzureichend erfasst werden. 17 Items werden somit eliminiert. Im Anschluss werden alle übrigen Items (mit $\lambda \geq .40$) unter inhaltlichen Aspekten und im Kontext ihres jeweiligen Faktors geprüft. Auf diese Weise können Items identifiziert werden, die inhaltlich nicht zum zu messenden Konstrukt (den übrigen Items) passen. Die Items werden erneut daraufhin überprüft, ob es eindeutig ist, welche epistemologische Merkmalsausprägung bei Zustimmen zum Item vorliegt. Außerdem werden bestehende Redundanzen zwischen Items identifiziert und aufgehoben, indem redundante Items mit niedrigerer Ladung und Kommunalität eliminiert werden.

**Faktor 1** wird dominiert von Items zu Überzeugungen zur Rechtfertigung des Wissens (RW) und Quelle des Wissens (QW), die nach Hofer und Pintrich den Bereich Überzeugungen zum Wissenserwerb (*nature of knowing*) bilden. Deutlich aus dem inhaltlichen Rahmen fallen die drei sw_21_u („Wenn man sich die Fakten aus einem pädagogischen Studientext aneignet, dann weiß man das Wichtigste."), gw_1_u („Die pädagogischen Studientexte enthalten stets gesichertes Wissen") und gw_7_u („Ich erwarte von meinen pädagogischen Studientexten gesichertes Wissen."). Item gw_4_u („Ich verlasse mich auf Forschungserkenntnisse in pädagogischen Studientexten, wenn zugehörige Studien genannt werden.") scheint zunächst aufgrund der Benennung (GW) auch inhaltlich aus dem Rahmen zu fallen, allerdings wird bei genauerer Betrachtung deutlich, dass es sich inhaltlich auch auf die Rechtfertigung des Wissens bezieht. Die Items qw_25_u („Ich zweifle nicht an, was in pädagogischen Studientexten steht.") und qw_20_u („Ich glaube generell schon, was in pädagogischen Studientexten steht.") sind sich inhaltlich sehr ähnlich und werden als redundant gewertet.

Das Item qw_20_u wird aufgrund niedrigerer Kommunalität zugunsten von qw_25_u eliminiert. Bei erneuter inhaltlicher Betrachtung von qw_10_u („Es ist besser, sich bei pädagogischen Studientexten an die großen Autoritäten im Fach zu halten.“) ist ungewiss, ob es eindeutig zwischen hohen und niedrigen Merkmalsausprägungen unterscheiden kann, d.h. ob Personen, die diesem Item zustimmen, wirklich notwendigerweise über absolutere epistemologische Überzeugungen verfügen. Item qw_23_u („Im Studium muss man sich darauf verlassen, dass das, was man in pädagogischen Studientexten liest, zutrifft.“) wird dadurch, dass sich alle Items aufgrund der Instruktion auf das pädagogische Studium beziehen, inhaltlich durch rw_33_u („Da ich nicht prüfen kann, ob stimmt was ich in pädagogischen Studientexten lese, glaube ich es erstmal.“) und qw_25_u („Ich zweifle nicht an, was in pädagogischen Studientexten steht.“) abgedeckt und entfällt deshalb. Die Items rw_30_u, sw_13_u, qw_6 und qw_31_u entfallen wegen zu niedriger Hauptladungen, wobei rw_30_u zusätzlich noch eine substanzielle Nebenladung aufweist. Qw_5, qw_8_u entfallen wegen zu hohen Nebenladungen (siehe Tab. 6.8, S. 198). Somit verbleiben in Dimension 1 die folgenden 9 Items: rw_33_u, qw_25_u, qw_21_u, qw_26_u, qw_13_u, rw_8_u, gw_4_u, rw_18_u, rw_29_u. Alle Items dieser Dimension sind gleichgerichtet formuliert. Eine hohe Zustimmung entspricht der Überzeugung, dass Wissen absolut und nicht relativ ist. Auf Basis der inhaltlichen Gemeinsamkeit der Items bildet dieser Faktor **Überzeugungen zur Absolutheit des Wissens** ab.

**Faktor 2** besteht überwiegend aus Items zur Struktur des Wissens (SW) und Gesichertheit des Wissens (GW), die nach Hofer und Pintrich (1997) den Bereich Überzeugungen zum Wissen (*nature of knowledge*) bilden. Qw_15_u und rw_32_u, qw_14_u und ew_11_u bilden hier inhaltlich die Ausreißer und werden aus diesem Grund entfernt. Die drei erstgenannten verfügen zudem über substanzielle Nebenladungen. Die drei Items sw_15_u („In pädagogischen Studientexten werden Dinge oft unnötig verkompliziert.“), sw_1_u („Die Dinge sind einfacher als sie in den meisten pädagogischen Studientexten dargestellt werden.“) und sw_27_u („Die meisten Autorinnen und Autoren von pädagogischen Studientexten stellen die Themen komplizierter dar, als sie eigentlich sind.“) sind inhaltlich redundant. Da sw_15_u über eine höhere Ladung und Kommunalität verfügt, werden sw_1_u und sw_27_u entfernt. Gw_3_u („Es ist Aufgabe der Expertinnen und Experten, die inhaltlichen Differenzen im Wissen auszudiskutieren. Erst danach sollten sie pädagogische Studientexte verfassen.“) und gw_2_u („Forschende sollten sich erstmal einig über ihre Aussagen werden, bevor sie diese in pädagogischen Studientexten anführen.“) sind sich inhaltlich so ähnlich, dass nur dasjenige Item mit der höheren Ladung (gw_3_u) beibehalten wird. Gw_11_u, sw_20_u und gw_10_u werden wegen zu niedriger Ladungen eliminiert. Gw_11_u verfügt zudem über eine Nebenladung (siehe Tab. 6.8, S. 198). Dimension 2 besteht nach der Itemreduktion aus den acht Items: sw_15_u, gw_13_u, gw_3_u, sw_2_u, sw_3_u, sw_11_u, sw_17_u und rw_36_u. Auch diese sind alle gleichgerichtet formuliert. Hier entspricht eine hohe Zustimmung der Überzeugung, dass Wissen einfach im Gegensatz zu komplex ist. Die Items des Faktors 2 erfassen die Dimension: **Überzeugungen zur Einfachheit des Wissens**.

**Faktor 3** enthält Items zu allen fünf intendierten Skalen: Quelle, Gesichertheit, Struktur, Rechterfertigung und Entwicklung des Wissens. Bei Analyse der Items wird jedoch deren inhaltliche Gemeinsamkeit ersichtlich. Die Items beziehen sich alle auf den Prozess des Wissenserwerbs. Gemeinsam liegt ihnen die Überzeugung zugrunde, dass der Wissenserwerb multimodal im Gegensatz zu unimodal zu erfolgen hat. Item rw_24 („Wenn ich einen pädagogischen Studientext lese, ist es mir wichtig, mir ein eigenes Erklärungsmodell der Inhalte zu bauen.“) wird aufgrund seiner inhaltlich Redundanz zu rw_27 („Es ist mir wichtig einen pädagogischen Studientext nicht nur zu lesen, sondern selbst zu durchdenken.“) -bei niedrigerer Ladung und Kommunalität- entfernt. Auch rw_11 („Es ist als Studierende/r wichtig, neben pädagogischen Studientexten noch weiteren Quellen zurate zu ziehen, um verschiedene Forschungsmeinungen zu konsultieren.“), qw_12 („Meine eigenen pädagogischen Ansichten erlange ich aus sehr vielen verschiedenen Quellen.“) und rw_26 („Den in einem pädagogischen Studientext aufgestellten Behauptungen kann man erst trauen, wenn man sich mit mehreren Wissensquellen zu dem Thema vertraut gemacht hat.“) sind sich inhaltlich sehr ähnlich, weshalb von diesen dreien nur Item rw_11 aufgrund dessen höherer Kommunalität beibehalten wird. Item rw_5 entfällt wegen einer substanziellen Doppelladung und Item rw_22 wegen einer zu geringen Hauptladung (siehe Tab. 6.9, S. 199). Dimension 3 umfasst somit nun noch die zehn Items qw_11, gw_9, sw_9, rw_27, gw_16, rw_12, rw_11, sw_22, ew_10 und gw_17. Wie bei den beiden vorherigen Dimensionen sind auch bei dieser alle Items gleichgerichtet formuliert. In diesem Fall entspricht eine hohe Zustimmung der Überzeugung, dass der Wissenserwerb multimodal ist oder sein sollte und nicht unimodal. Die Items des Faktors 3 erfassen die Dimension: **Überzeugungen zur Multimodalität des Wissenserwerbs**.

In **Faktor 4** dominieren ganz deutlich die Items zu den Überzeugungen zur Entwicklung des Wissens. Item sw_19_u („Wirklich relevantes Wissen ist in pädagogischen Studientexten immer widerspruchsfrei.“) fällt dabei inhaltlich deutlich aus dem Rahmen und wird deshalb entfernt. Item ew_5_u wird aufgrund seiner niedrigen Ladung eliminiert. Die Items ew_5_u („Bei pädagogischen Studientexten kommen über die Jahre zwar ein paar Dinge hinzu aber dennoch bleiben die Erkenntnisse weitestgehend erhalten.“) und ew_9_u („Heutzutage finden Forschende eher nur sehr spezielle Kleinigkeiten an neuem pädagogischen Wissen heraus, deshalb bleiben die Inhalte von Studientexten im Großen und Ganzen über die Jahrzehnte gleich.“) sind sich inhaltlich ähnlich. Das zweite wird aufgrund der ungünstigen Formulierung und der geringeren Ladung und Kommunalität eliminiert. Ew_8, ew_3 und ew_5_u verfügen über Ladungen von .35 bis .39, wobei letzteres noch eine Nebenladung aufweist. Insgesamt ist der vierte Faktor nicht sonderlich stark, allerdings wird der inhaltliche Aspekt der Wissensentwicklung von keiner der anderen Dimensionen abgedeckt. Vor der Eliminierung eines jeden Items mit niedrigen Ladungen sollte immer die Inhaltsvalidität berücksichtigt werden. Ist das Item für das Konstrukt wichtig, so empfiehlt Bühner (2011, S. 350), es trotz geringer Ladung zunächst beizubehalten. Item ew_5_u („Bei pädagogischen Studientexten kommen über die Jahre zwar ein paar Dinge hinzu aber dennoch bleiben die Erkenntnisse weitestgehend erhalten.“) ist ähnlich zu ew_16_u („Was heute

in pädagogischen Studientexten als wahr gilt, trifft in Zukunft auch noch zu.“), unterscheidet sich jedoch in seiner psychometrischen Schwierigkeit und wird deshalb beibehalten. Auch das Item ew_8 („Ich werde mich schon in wenigen Jahren den meisten Inhalten aus den pädagogischen Studientexten neu widmen müssen, weil viele Dinge wieder überholt sein werden.“) wird beibehalten, da es einen Aspekt der Entwicklung des Wissens abdeckt, der in keinem der anderen Items bisher berücksichtigt wird. Item ew_3 („Wie etwas verstanden wird, kann sich in der pädagogischen Forschung sehr schnell ändern. Demnach ändert sich dann auch, was in Studientexten berichtet wird.“) wird aufgrund der relativ guten Kommunalität von .32 vorerst beibehalten (siehe Tab. 6.9, S. 199). Aus diesem Grund wird vorgezogen, eher mehr als weniger Items für die zweite Stichprobe beizubehalten. Die vierte Dimension umfasst die sieben Items ew_13, ew_12, ew_2, ew_16_u, ew_5_u, ew_3, ew_8. Fünf der Items sind gleichgerichtet formuliert. Item ew_5_u und ew_16_u sind im Vergleich zu den anderen Items dieser Dimension inhaltlich umgekehrt formuliert, d.h. eine hohe Zustimmung drückt die Überzeugung aus, dass Wissen eher statisch ist und sich nicht entwickelt, während eine hohe Zustimmung bei den anderen Items die Überzeugung ausdrückt, dass Wissen sich weiterentwickelt. Dieser Faktor erfasst **Überzeugungen zur Entwicklung des Wissens**.

**Faktor 5** besteht aus nur fünf Items zur Rechtfertigung des Wissens (RW), von denen allerdings nur eines (rw_2_u) eine Ladung von $\lambda >\mid .40 \mid$ und kein einziges eine hohe Ladung ($\lambda > .60$) aufweist. Zwei der Items, rw_37 und rw_13, verfügen zudem über substanzielle Nebenladungen, bei gleichzeitig relativ geringen Hauptladungen von .39 und .33 (siehe Tab. 6.9, S. 199). Da dieser Faktor so schwach repräsentiert wird (vgl. Kap. 4.3.1, S. 153) und der inhaltliche Aspekt Rechtfertigung des Wissens des Konstrukts der epistemologischen Überzeugungen inhaltlich im ersten Faktor bereit abgedeckt wird, entfällt dieser Faktor mit seinen fünf Items ganz.

### 6.3.5 Prüfung der Faktorstruktur nach Itemreduktion

Mit dem Entfernen von Items kann es vorkommen, dass die Faktorstruktur sich ändert (vgl. Bühner, 2011; Brown, 2015, S. 34). Aus diesem Grund muss sie bei der 34-Item Version des StEB in einer erneuten EFA auf Stabilität hin überprüft werden (vgl. Kap. 4.3.1, S.153). Wie zuvor auch, wird mit dem bereinigten Datensatz mit imputierten fehlenden Werten gearbeitet, bei dem die umgekehrt formulierten Items umkodiert sind.

Das itemübergreifende KMO-Maß bescheinigt nach der Itemreduktion mit KMO = .80, gute Eignung der Daten für die Durchführung der Faktorenanalyse. Auch die itemspezifischen MSA-Maße sind mit .67 bis .86 ausnahmslos annehmbar (siehe Tab. 4.2, S. 157). In Bezug auf die optimale Anzahl an zu extrahierenden Faktoren kommen sowohl der Scree-Test, AC, OC als auch die Parallelanalyse zu einer Faktorenanzahl von vier (siehe Abb. 6.3).

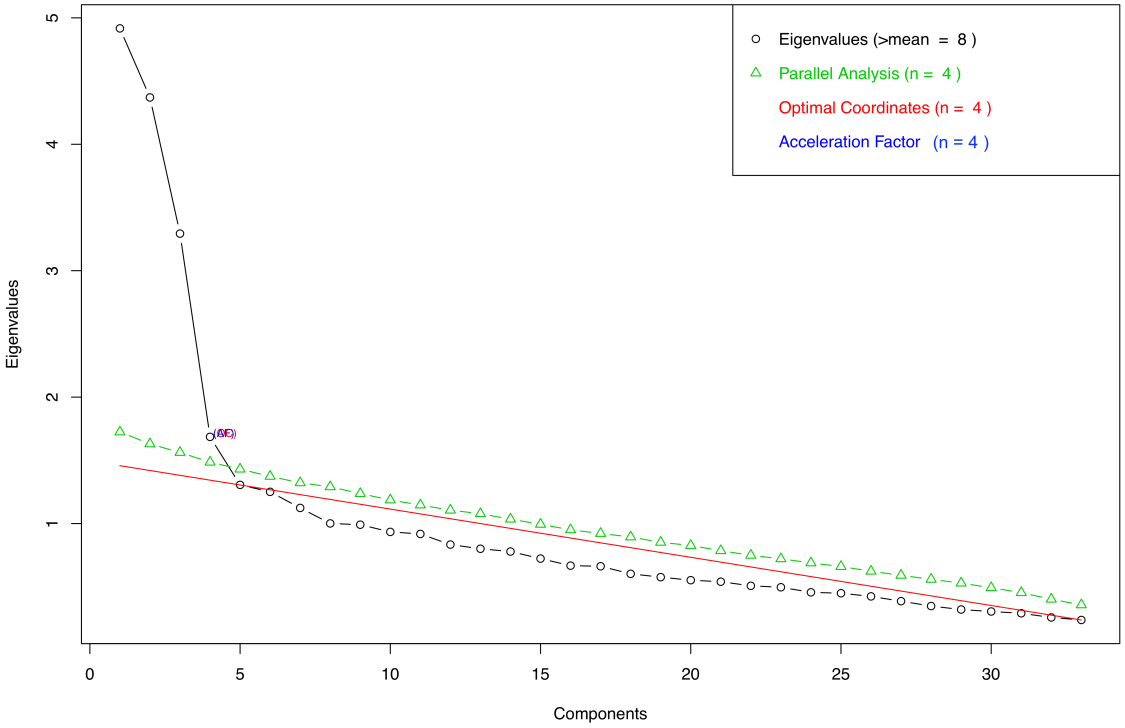


ABBILDUNG 6.3: Kombinierte Darstellung verschiedener Tests zur Bestimmung der geeigneten Faktorenanzahl

In diesem Sinne wurde eine Hauptachsenanalyse mit Promax-Rotation unter Extraktion von vier Faktoren über die 34 verbleibenden Überzeugungsitems berechnet. Die Musterladungsmatrix, Kommunalitäten und Uniqueness sind in Tabelle 6.10 (S. 205) dargestellt. Ladungen unter $\lambda = | .20 |$ werden in der Tabelle nicht aufgeführt. Anhand der Mustermatrix wird deutlich, dass die Faktorstruktur über die Itemreduktion hinweg stabil geblieben ist. Alle 34 verbleibenden Items laden weiterhin auf den im Rahmen der ersten EFA (siehe Tabellen 6.8 und 6.9 S. 198, 199) identifizierten Faktoren. Alle Items, außer die Items ew_5_u ($\lambda$ = .27), laden mit $\lambda \geq | .32 |$ substantiell auf ihrem jeweiligen Faktor. Es kommen zwar ein paar Nebenladungen vor, diese sind allerdings angemessen gering ($\lambda < | .32 |$). In allen vier

Faktoren kommen einige Kernitems mit hohen Faktorladungen ($\lambda >| .60 |$) und wenigstens mittleren Kommunalitäten ($h^2$ = .40 - .60) vor.

TABELLE 6.10: Mustermatrix mit Faktorladungen $\lambda$ und zugehörigen Kommunalitäten und Uniqueness der 34-Item 4-Faktorlösung ($N$ = 217)

| Items | Faktor 1 | Faktor 2 | Faktor 3 | Faktor 4 | $h^2$ | $u^2$ |
|---|---|---|---|---|---|---|
| rw_33_u | .79 | | | | .51 | .49 |
| qw_25_u | .69 | | | | .49 | .51 |
| qw_13_u | .63 | | | | .43 | .57 |
| qw_21_u | .60 | | | | .39 | .61 |
| qw_26_u | .58 | | | | .32 | .68 |
| rw_29_u | .56 | | | | .34 | .66 |
| gw_4_u | .55 | | | | .31 | .69 |
| rw_8_u | .51 | | | | .31 | .69 |
| rw_18_u | .50 | | | | .37 | .63 |
| sw_15_u | -.25 | .73 | | | .57 | .43 |
| sw_2_u | | .72 | | | .55 | .45 |
| sw_17_u | .28 | .61 | | | .52 | .48 |
| gw_3_u | | .57 | | | .31 | .69 |
| sw_11_u | | .55 | | | .30 | .70 |
| gw_13_u | | .54 | | | .34 | .66 |
| rw_36_u | | .41 | | | .29 | .71 |
| sw_3_u | | .38 | | | .21 | .79 |
| sw_9 | | | .67 | | .43 | .57 |
| qw_11 | | .26 | .66 | | .57 | .43 |
| rw_11 | .26 | | .62 | | .40 | .60 |
| rw_27 | | | .57 | | .33 | .67 |
| rw_12 | -.20 | | .57 | | .41 | .59 |
| gw_17 | | | .55 | | .31 | .69 |
| gw_16 | | | .52 | | .24 | .76 |
| gw_9 | | | .49 | | .25 | .75 |
| sw_22 | | | .46 | | .22 | .78 |
| ew_10 | | | .43 | | .23 | .77 |
| ew_13 | | | | .79 | .52 | .48 |
| ew_12 | | | | .75 | .45 | .55 |
| ew_3 | | | | .56 | .40 | .60 |
| ew_2 | | | | .40 | .23 | .77 |
| ew_8 | | .21 | | .40 | .27 | .73 |
| ew_16_u | .26 | | | .34 | .25 | .75 |
| ew_5_u | | | .20 | .27 | .19 | .81 |

Wie schon bei den Ergebnissen der ersten EFA zeigt sich allerdings, dass mehr Items niedrige Kommunalitäten ($h^2$ = .20 - .39) als mittlere erreichen. Item ew_5_u („Bei pädagogischen Studientexten kommen über die Jahre zwar ein paar Dinge hinzu aber dennoch bleiben die Erkenntnisse weitestgehend erhalten.“) erreicht als einziges Item nur eine sehr niedrige Kommunalität von $h^2$ < .20. Es zeigt sich hier, dass Item ew_5_u deutliche Schwächen aufweist. Da diese Ergebnisse jedoch nicht der erneuten Itemreduktion dienen, die im vorherigen Schritt vorgenommen wurde, sondern lediglich die Stabilität der Faktorstruktur geprüft werden soll, wird Item ew_5_u trotz seiner geringen Ladung und Kommunalität zunächst beibehalten, um es im Rahmen der Validierungsstichprobe erneut zu prüfen.

## 6.4 Psychometrische Eigenschaften der Subskalen

Die Prüfung der psychometrischen Eigenschaften einer Testskala umfasst nach der klassischen Testtheorie eine Item- und Skalenanalyse (Bühner, 2011). Im Rahmen der Itemanalyse werden die StEB-Items zunächst im Hinblick auf ihre psychometrische Schwierigkeit und Trennschärfe untersucht. Im Anschluss wird die interne Konsistenz der Skalen und deren Homogenität untersucht. Die Item- und Skalenanalyse erfolgt hier weniger ausführlich als im Rahmen der Analyse des fertigen StEB Inventars (siehe Kap. 7.6, S. 263), da es sich bei dieser Version des Instruments noch nicht um die abschließende Version, sondern lediglich um einen Zwischenschritt in der Instrumententwicklung handelt.

### 6.4.1 Itemanalyse

„Die psychometrische Itemschwierigkeit ist als Zustimmung zu einem Item in Schlüsselrichtung der Skala definiert." (Bühner, 2011, S. 219). Wie im Zusammenhang mit der deskriptiven Analyse der StEB-Items (siehe Abschnitt 6.2.1, S. 188) bereits kurz angeschnitten, ist die Analyse der Itemschwierigkeit für die Instrumententwicklung interessant, da die konstruierten Items darauf abzielen, unterschiedliche Merkmalsausprägungen zwischen Versuchspersonen auch als solche zu erfassen (Moosbrugger, 2012, S. 76). Items, die zu leicht oder zu schwierig sind, können Merkmalsdifferenzen zwischen Personen nicht angemessen erfassen. Für die Interpretation der Itemmittelwerte als Maß für die Itemschwierigkeit ist dabei bedeutsam, wie viele Antwortstufen das Antwortformat aufweist (Bühner, 2011, S. 219 f.) und ob es umkodiert wurde. Um die Interpretation der jeweiligen Itemschwierigkeiten zu erleichtern, werden die Items vor der Berechnung des Schwierigkeitsindex nicht umkodiert. Die Itemmittelwerte für die Gesamtstichprobe wurden bereits im Zusammenhang mit der deskriptiven Analyse der StEB-Items (siehe Abschnitt 6.2.1, S. 188) berichtet. Dort wurden allerdings die umgekehrt formulierten Items vorher umkodiert. Aus diesem Grund werden die Itemmittelwerte und zugehörige Standardabweichungen erneut berichtet. Die Mittelwerte und Standardabweichungen basieren ausschließlich auf den realen Itemantworten der befragten Personen und enthalten keine imputierten Werte.

Aus Tabelle 6.11 auf Seite 208 gehen die Itemmittelwerte und Standardabweichungen hervor. Das Item mit dem höchsten Mittelwert ($M$ = 4.08 und $SD$ = 0.67) ist gw_17 („Ich bin offen gegenüber Ansichten in pädagogischen Studientexten, die ich bisher noch nicht kannte."). Der für eine fünfstufige Skala hohe Mittelwert zeigt, dass viele Befragte dieser Aussage des Items zustimmen. Dem umgekehrt formulierten Item sw_3_u („Ich bin der Überzeugung, dass Autorinnen und Autoren, die die Sachverhalte in pädagogischen Studientexten komplex darstellen, sich eigentlich gar nicht so gut auskennen.") wird dagegen mit einem Mittelwert von 2.10 ($SD$ = .83) von allen Items durchschnittlich am wenigsten zugestimmt. Im Hinblick auf die fünf möglichen Antwortstufen der Items sind die gefundenen Itemschwierigkeiten nicht als zu extrem einzuschätzen. Die Standardabweichungen (siehe Tab. 6.11, S. 208) geben Hinweise darauf, wie stark die Itemantworten der Individuen vom Mittelwert abweichen. Item gw_17 verfügt gleichzeitig über die geringste Standardabweichung. Die Itemantworten bewegen sich also durchschnittlich sehr nah um den

Itemmittelwert.

Außerdem werden in Tabelle 6.11 (S. 208) der itemspezifische Schwierigkeitsindex $P_i$ nach Moosbrugger (2012, S. 76) sowie die part-whole korrigierte Itemtrennschärfe $r_{ix}$ berichtet (siehe Tabelle 6.11), die im Folgenden erläutert werden. Neben dem Mittelwert als Maß für die Itemschwierigkeit lässt sich der Schierigkeitsindex $P_i$ nach Moosbrugger (2012, S. 77) berechnen. Beide müssten zum gleichen Ergebnis kommen. $P_i$ ergibt sich aus dem Quotient aus der für ein Item real erreichten Punktesumme aller Befragten und der maximal erreichbaren Punktsumme, multipliziert mit 100. $P_i$ bewegt sich somit zwischen 0 und 100. Je mehr Befragte „symptomatisch im Sinne des zu erhebenden Merkmals“ (Moosbrugger, 2012, S. 77) antworten, desto größer wird $P_i$. Items mit hohen $P_i$-Werten sind demnach leichter als Items mit niedrigen $P_i$-Werten. Somit könnte man $P_i$ auch als Leichtigkeitsindex betrachten. Zu berücksichtigen ist im vorliegenden Fall, dass zwei der Skalen ausschließlich aus umgekehrt formulierten Items bestehen. Bei den Skalen Einfachheit und Absolutheit drückt eine hohe Zustimmung aus, dass Wissen als absoluter bzw. als einfacher betrachtet wird. Bei der Berechnung des Schwierigkeitsindex $P_i$ wurde berückichtigt, dass der Wertebereich der Ratingskala von 1 bis 5 reicht und nicht etwa mit 0 beginnt, d.h. dass auch bei vollständiger Ablehnung eines Items noch eine Mindestpunktzahl von 1 erreicht wird. Auch die berechneten Schwierigkeitsindizes (siehe Tabelle 6.11, S. 208) basieren ausschließlich auf den realen Itemantworten der befragten Personen und enthalten keine imputierten Werte. sw_3_u ist mit $P_i$ = 27.43 am psychometrisch schwierigsten, während gw_17 mit $P_i$ = 76.96 am psychometrisch leichtesten ist – übereinstimmend oben berichteten Mittelwerten als Maß der Itemschwierigkeit.

Die Itemtrennschärfe $r_{ix}$ gibt an, wie gut das einzelne Item das Gesamtergebnis eines Tests repräsentiert. Sie muss demnach für jedes Item einer Skala getrennt berechnet werden. Sie „[...] ist definiert als die Korrelation der Beantwortung dieses Items mit dem Gesamttestwert.“' (Döring & Bortz, 2016, S. 476 ff.). Da für das StEB Inventar eine mehrdimensionale Struktur angenommen wird, die im Rahmen der faktorenanalytischen Auswertung dieser ersten Stichprobe gezeigt werden konnte, muss zur Berechnung der Trennschärfe statt des Gesamttestwerts der Gesamtwert einer Skala verwendet werden. Statt der Trennschärfe wird normalerweise – und so auch hier – die korrigierte Trennschärfe $r_{ix}$ für ein Item bestimmt. Dabei wird der Wert des Items, für das die Trennschärfe bestimmt wird, bei der Berechnung des Gesamttestwertes nicht berücksichtigt (*part-whole Korrektur*). Das hat den Grund, dass bei dem herkömmlichen Verfahren, bei dem der Wert des Items in den Gesamttestwert miteinfließt, die Korrelation künstlich in die Höhe getrieben, das heißt verfälscht wird (Fisseni, 1990, S. 40 f.). Ein hoher Trennschärfewert von über .50 wird angestrebt. Ein Wert von .30 bis .50 ist als mittelmäßige Trennschärfe einzustufen (Döring & Bortz, 2016, S. 476 ff.). Der Trennschärfewert ist folgendermaßen zu interpretieren: Ein hoher Trennschärfewert gibt an, dass eine Person, die im Gesamtergebnis einer Skala einen hohen Wert erreicht, in einem trennscharfen Item dieser Skala ebenfalls einen hohen Wert erreicht. So lässt sich aus einem trennscharfen Einzelitem das Gesamtergebnis in dieser Skala vorhersagen. Dieses Gesamtergebnis der Skala steht für die Ausprägung des gemessenen Merkmals.

TABELLE 6.11: Itemschwierigkeiten $P_i$ und korrigierte Itemtrennschärfen $r_{ix}$ der 34-Item 4-Faktorlösung

| Items | $P_i$ | $M$ | $SD$ | $r_{ix}$ |
|---|---|---|---|---|
| Absolutheit | | | | |
| rw_33_u | 47.00 | 2.88 | 1.02 | .64 |
| qw_25_u | 37.10 | 2.48 | .98 | .59 |
| qw_13_u | 41.13 | 2.65 | .89 | .56 |
| qw_21_u | 51.29 | 3.05 | .78 | .55 |
| qw_26_u | 45.28 | 2.81 | 1.16 | .51 |
| rw_29_u | 50.46 | 3.02 | 1.00 | .53 |
| gw_4_u | 60.76 | 3.43 | .83 | .44 |
| rw_8_u | 52.88 | 3.12 | .91 | .45 |
| rw_18_u | 40.97 | 2.64 | .99 | .53 |
| Einfachheit | | | | |
| sw_15_u | 69.21 | 3.77 | 1.03 | .60 |
| sw_2_u | 63.77 | 3.55 | .96 | .63 |
| sw_17_u | 61.41 | 3.46 | .94 | .59 |
| gw_3_u | 56.60 | 3.26 | 1.05 | .47 |
| sw_11_u | 60.83 | 3.43 | .98 | .49 |
| gw_13_u | 47.81 | 2.91 | 1.15 | .50 |
| rw_36_u | 47.58 | 2.90 | 1.14 | .44 |
| sw_3_u | 27.43 | 2.10 | .83 | .37 |
| Multimodalität | | | | |
| sw_9 | 67.13 | 3.69 | .89 | .53 |
| qw_11 | 64.40 | 3.58 | 1.03 | .64 |
| rw_11 | 74.31 | 3.97 | .83 | .50 |
| rw_27 | 74.07 | 3.96 | .98 | .51 |
| rw_12 | 63.67 | 3.55 | .94 | .49 |
| gw_17 | 76.96 | 4.08 | .67 | .48 |
| gw_16 | 60.25 | 3.41 | .89 | .42 |
| gw_9 | 68.40 | 3.74 | .91 | .44 |
| sw_22 | 74.65 | 3.99 | .80 | .42 |
| ew_10 | 62.67 | 3.51 | .87 | .40 |
| Entwicklung | | | | |
| ew_13 | 75.82 | 4.03 | .74 | .56 |
| ew_12 | 70.51 | 3.82 | .90 | .52 |
| ew_3 | 64.81 | 3.59 | .85 | .51 |
| ew_2 | 69.82 | 3.79 | .86 | .32 |
| ew_8 | 60.37 | 3.41 | .82 | .40 |
| ew_16_u | 38.31 | 2.53 | .76 | .38 |
| ew_5_u | 47.58 | 2.90 | .83 | .25 |

Somit lässt sich also aus einem trennscharfen Item bereits vorhersagen, ob die jeweiligen Versuchspersonen über eine hohe oder niedrige Ausprägung des gemessenen Konstrukts verfügen. Auf diese Art trennt das Item erfolgreich die Gruppe mit niedriger Merkmalsausprägung von der Gruppe mit der hohen Merkmalsausprägung (Döring & Bortz, 2016, S. 476 ff.).

Tabelle 6.11 auf Seite 208 ist zu entnehmen, dass die Items der vier Subskalen, mit Ausnahme eines Items, mit $r_{ix}$ von .38 bis .64 über eine mittlere bis hohe Trennschärfe verfügen. Item ew_5_u bildet hier mit einer sehr geringen korrigierten Trennschärfe von .25 die Ausnahme. Das bedeutet, dass dieses Item als ein schlechter Indikator des anstrebten Konstrukts Überzeugungen zur Entwicklung des Wissens anzusehen ist, da es Informationen generiert, die nicht mit dem Gesamttestergebnis übereinstimmen.

### 6.4.2 Skalenanalyse

Im Rahmen der Skalenanalyse wird die Skalenhomogenität sowie die interne Konsistenz Cronbachs Alpha für jede der vier Skalen bestimmt. Ein mehrdimensionaler Test wie das konstruierte StEB Inventar besteht aus verschiedenen Skalen, die jeweils ein unterschiedliches latentes Merkmal messen. Die Items einer Skala dienen der Operationalisierung des zu messenden Konstrukts. Da alle Items einer Skala gemeinsam auf ein bestimmtes Konstrukt abzielen, müssen die Items untereinander in Zusammenhang stehen. Die Höhe dieses geforderten Zusammenhangs nennt man Homogenität. Die Itemhomogenität wird aus den Iteminterkorrelationen berechnet, indem der Durchschnitt der Korrelation eines Items mit allen anderen Items gebildet wird. Mittelt man die Homogenität aller Items (*Mean-Inter-Item-Korrelation*, MIC) einer Skala ergibt sich die Skalenhomogenität (Döring & Bortz, 2016, S. 480 f.), die in Tabelle 6.12 zu finden sind.

TABELLE 6.12: Skalenhomogenität und interne Konsistenz der Skalen

| | Absolutheit | Einfachheit | Multimodalität | Entwicklung |
|---|---|---|---|---|
| $MIC$ | .35 | .33 | .29 | .26 |
| $\alpha$ | .83 | .80 | .81 | .71 |

Mit Werten zwischen .26 und .35 liegen die Skalenhomogenitäten innerhalb des von Briggs und Cheek (1986, S. 115) empfohlenen Akzeptanzbereichs für die Skalenhomogenität (.20 bis .40). Eine Skalenhomogenität ab einem Wert von .40 ist als zu hoch einzustufen, da höhere Homogenitäten bedeuten, dass die Skalen sich inhaltlich zu ähnlich seien. Im Rahmen dieser Instrumententwicklung ist zu berücksichtigen, dass das StEB Inventar über eine Hauptdimension „epistemologische Überzeugungen zum erziehungswissenschaftlichen Wissen beim Lernen mit Studientexten“ verfügt. Das bedeutet, dass die verschiedenen Skalen vier unterschiedliche Aspekte eines übergeordneten Konstrukts messen. Das heißt, mit einem gewissen Zusammenhang zwischen den Skalen ist nicht nur zu rechnen, sondern dieser ist erwünscht, denn wenn alle Skalen völlig unabhängig voneinander wären, könnten sie nicht gemeinsam ein Hauptkonstrukt erheben.

Die im Rahmen der EFA ermittelten Zusammenhänge zwischen den vier Faktoren der 34-Item StEB-Version sind in Tabelle 6.13 einzusehen. Hier ist bei der Interpretation zu berücksichtigen, dass die Items für die EFA alle so kodiert waren, dass eine höhere Zustimmung für höherere Sophistiziertheit in den wissensbezogenen Überzeugungen steht. Je geringer der Zusammenhang ist, desto unähnlicher sind sich die Skalen, je höher, desto ähnlicher (Bühner, 2011, S. 80).

TABELLE 6.13: Korrelationen zwischen den Faktoren nach dem Ergebnis der EFA der 34-Item StEB-Version

| Subskalen | Absolutheit | Einfachheit | Multimodalität |
|---|---|---|---|
| Absolutheit | - | | |
| Einfachheit | .18 | - | |
| Multimodalität | -.13 | .21 | - |
| Entwicklung | .38 | -.06 | .28 |

Es bestehen positive Zusammenhänge zwischen dem Faktor *Absolutheit des Wissens* mit den Faktoren *Einfachheit des Wissens* und *Entwicklung des Wissens*, jedoch ein negativer Zusammenhang mit dem Faktor *Multimodalität des Wissenserwerbs*. Es besteht kein Zusammenhang zwischen den Dimensionen *Einfachheit des Wissens* und *Entwicklung des Wissens*, allerdings ein positiver Zusammenhang zwischen *Einfachheit des Wissens* und *Multimodalität des Wissenserwerbs*. Die Faktoren *Multimodalität des Wissenserwerbs* und *Entwicklung des Wissens* korrelieren positiv.

Die Überzeugung, dass Wissen aus erziehungswissenschaftlichen Studientexten relativ (und nicht absolut) ist, korreliert negativ mit der Überzeugung, dass der erziehungswissenschaftliche Wissenserwerb multimodal sein muss (und nicht unimodal). Die Überzeugung, dass Wissen relativ ist, korreliert positiv mit der Überzeugung, dass Wissen komplex strukturiert ist. Die Überzeugung, dass Wissen relativ ist geht mit der Überzeugung einher, dass Wissen sich entwickelt. Während die Überzeugung, dass Wissen komplex strukturiert ist, positiv mit der Überzeugung zusammenhängt, Wissen werde multimodal erworben. Die Überzeugung, dass erziehungswissenschaftliches Wissen komplex strukturiert ist, korreliert dagegen nicht (gegen 0) mit der Überzeugung, dass sich erziehungswissenschaftliches Wissen entwickelt.

Wie erwartet bestehen zwischen den vier Dimensionen epistemologischer Überzeugungen zum erziehungswissenschaftlichen Wissen tatsächlich Zusammenhänge. Die Stärke dieser Zusammenhänge ist jedoch nicht übermäßig groß, was darauf hindeutet, dass die vier Subskalen des StEB tatsächlich unterschiedliche Teilaspekte des Gesamtkonstrukts erfassen und nicht etwa die gleichen Aspekte (Bühner, 2011, S. 80).

## 6.5 Diskussion und Rückbezug zum theoretischen Hintergrund

Die empirische Prüfung des 98-Item StEB Testentwurfs umfasste den Einsatz an einer Stichprobe von zunächst 221 Lehramtsstudierenden zwei deutscher Universitäten. Nach der Bereinigung der Daten verblieben 217 Datensätze, von denen 187 bezüglich der Überzeugungsitems vollständig waren. Die Analyse fehlender Werte ergab eine insgesamt geringe Anzahl fehlender Werte und darüber hinaus, dass den fehlenden Werten keine Systematik zugrunde liegt (MCAR), was den Einsatz der *arithmetic mean imputation* als Methode zum Umgang mit den fehlenden Werten erlaubte. Die fehlenden StEB-Itemwerte wurden durch den Stichprobenmittelwert der jeweils betroffenen Variable ersetzt. Neun Überzeugungsitems wurden daraufhin von den weiteren Analysen ausgeschlossen, da sich im Rahmen der Prüfung der Verteilungen herausstellte, dass die Befragten in ihrer Itembeantwortung nicht die volle Breite des Antwortspektrums abdeckten. Der Analyse der Itemverteilungen schloss sich die Prüfung der Voraussetzungen für den Einsatz der EFA an. Hier wurde analysiert, ob die Stichprobengröße ausreichend ist und mittels KMO-Kriterium und itemspezifischem MSA untersucht, ob sich die Itemkorrelationsmatrix für die EFA eignet. Dies hatte den Ausschluss von drei weiteren Items aufgrund zu geringer MSA-Werte zur Folge. Im Anschluss wurde die Hauptachsenanalyse als geeignetes verteilungsfreies Extraktionsverfahren gewählt, da der Mardia Test ergab, dass für die Überzeugungsitems keine multivariate Normalverteilung angenommen werden kann und sich dieses Verfahren auch speziell für Ratingitems eignet. Zur anschließenden Bestimmung der geeigneten Faktorenanzahl für die EFA wurden neben den inhaltlich-theoretischen Überlegungen die Ergebnisse der Parallelanalyse, des Scree-Tests, der MAP-Methode sowie der OC- und AF-Methode berücksichtigt. Ein Vergleich der Ergebnisse der fünf berechneten EFAs (mit obliquer Rotation) ergab, dass die Extraktion von fünf Faktoren zu einem inhaltlich-theoretisch interpretierbaren Ergebnis führte. Die resultierende Mustermatrix wurde bezüglich der Höhe der Ladungen und Kommunalitäten untersucht, die neben dem Iteminhalt eine bedeutende Rolle bei der Itemreduktion spielten. Nach der Itemreduktion schloss sich eine erneute faktorenanalytische Prüfung der Faktorstruktur an, die ergab, dass die Faktorstruktur der verbliebenen 34-Item StEB-Version stabil bleibt. Für die 34-Item StEB-Version wurden im Rahmen einer Item- und Skalenanalyse die psychometrischen Itemschwierigkeiten und Itemtrennschärfen sowie die Skalenhomogenitäten und -reliabilitäten bestimmt.

Die aus der ersten Studie resultierende 34-Item StEB Inventar Version verfügt über die vier Dimensionen der Überzeugungen zur *Absolutheit des Wissens*, *Einfachheit des Wissens*, *Multimodalität des Wissenserwerbs* und *Entwicklung des Wissens*. Aus Tabelle 6.14 auf Seite 212 ist die Anzahl der Items je Faktor zu entnehmen, die Anzahl der invertiert gepolten Items sowie die Anzahl der auf bestehenden Fragebögen basierenden Items (abgewandelt) je Skala. Hier ist auch zu entnehmen, welche epistemologische Überzeugungsausprägung einer Zustimmung bzw. Ablehung zu den Items der jeweilige Subskala entspricht. Wie oben bereits erwähnt, verfügen drei der vier Subskalen über durchweg gleichgerichtet formulierte Items. Über alle Items der Skalen Absolutheit und Einfachheit hinweg gilt zunächst (sofern

diese nicht umgepolt werden), je höher die Zustimmung desto *weniger* sophistiziert ist die epistemologische Sichtweise auf Wissen. Bei der Skala Multimodalität verhält es sich anders herum, höhere Zustimmung stellt einheitlich für alle Items dieser Skala eine sophistiziertere wissensbezogene Überzeugungsausprägung dar. Die Skala Entwicklung des Wissens verfügt über zwei invertiert gepolte Items, während die übrigen fünf Items so formuliert sind, dass Zustimmung einer sophistizierten Überzeugung zur Entwicklung des Wissens entspricht. Die Merkmalsausprägung, die inhaltlich einer Zustimmung zu den Items entspricht (rechter Pol), wurde für die Benennung der Subskalen verwendet.

TABELLE 6.14: Item-Skalen-Zuordnung des 34-Item StEB Inventars

| Dimension | Item-anzahl | Itemanzahl invertiert | Itemanzahl abgewandelt | Ablehnung linker Pol | Zustimmung rechter Pol |
|---|---|---|---|---|---|
| Absolutheit des Wissens | 9 | 9 | 2 | relativ | absolut |
| Multimodalität des Wissenserwerbs | 10 | 0 | 1 | unimodal | multimodal |
| Einfachheit des Wissens | 8 | 8 | 0 | komplex | einfach |
| Entwicklung des Wissens | 7 | 2 | 3 | statisch | entwickelnd |
| Gesamt | 34 | 19 | 6 | | |

Der Wortlaut der 34 Items ist im Anhang B in den Tabellen B.14 und B.15 (ab S. 382) einsehbar. Ebenfalls im Anhang B in Tabelle B.18 (S. 386) sind die sechs Items aufgeführt, die in ihrer Formulierung auf Items aus bestehenden Fragebögen basieren. Der Itemwortlaut dieser ursprünglichen Items ist dort ebenfalls einzusehen.

Bei der Itemformulierung in dieser Testentwicklung (siehe Kap. 5.2, ab S. 174) wurde eine Mischung aus negativ und positiv gepolten Items entwickelt, um bei der Itembeantwortung Akquieszenz, d.h. einen zustimmenden Antwortstil, vorzubeugen (vgl. Bühner, 2011, S. 134). Dazu wurden Items, bei denen eine hohe Antwortkategorie einer starken Merkmalsausprägung – einer sophistizierten wissensbezogenen Überzeugung – entspricht, mit solchen vermischt, die anders herum aufgebaut sind. Bei einer solchen Mischung von negativ und positiv gepolten Items besteht allerdings die Gefahr, dass sich diese Polung auf die Faktorstruktur auswirkt und sich die negativ und positiv gepolten Items getrennt zu Skalen clustern (Bühner, 2011, S. 134). Man spricht in so einem Fall, wenn also die Polung für die Faktorstruktur ursächlich ist und nicht unbedingt die zugrundeliegende latente Variable, von einem Methodeneffekt (Brown, 2015; Bühner, 2011). Da im vorliegenden Fall tatsächlich Skalen mit einheitlicher Polung entstanden sind, könnte dies den Eindruck eines solchen Methodeneffekts erwecken. Wie im Folgenden jedoch dargestellt wird, sind die Zusammenstellungen der Items, d.h. die faktorenanalytisch gebildeten Skalen epistemologischer Überzeugungen zum erziehungswissenschaftlichen Wissen inhaltlich sinnvoll

interpretierbar. Die Clusterung der Items ist also im Hinblick auf die theoretische Konzeptualisierung von Hofer und Pintrich (1997) und Hofer (2001) klar erklärbar und stützt die theoretische Konzeption. Dies entkräftigt die Annahme eines Methodeneffekts.

Zentral für diese Arbeit ist die Fragestellung, ob sich die epistemologischen Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen aus Studientexten in der Faktorstruktur des StEB, konform zum Theoriegefüge abbilden lassen (vgl. Kap. 3, S. 135). Daraus ergeben sich die Fragen, welche inhaltlichen Faktoren sich für das StEB Inventar identifizieren lassen und ob diese resultierenden Dimensionen mit dem zugrundegelegten Theoriegefüge epistemologischer Überzeugungen vereinbar sind. In Abbildung 6.4 wird die Gegenüberstellung der theoretisch angenommenen und der empirisch gefunden Faktorstruktur gezeigt.

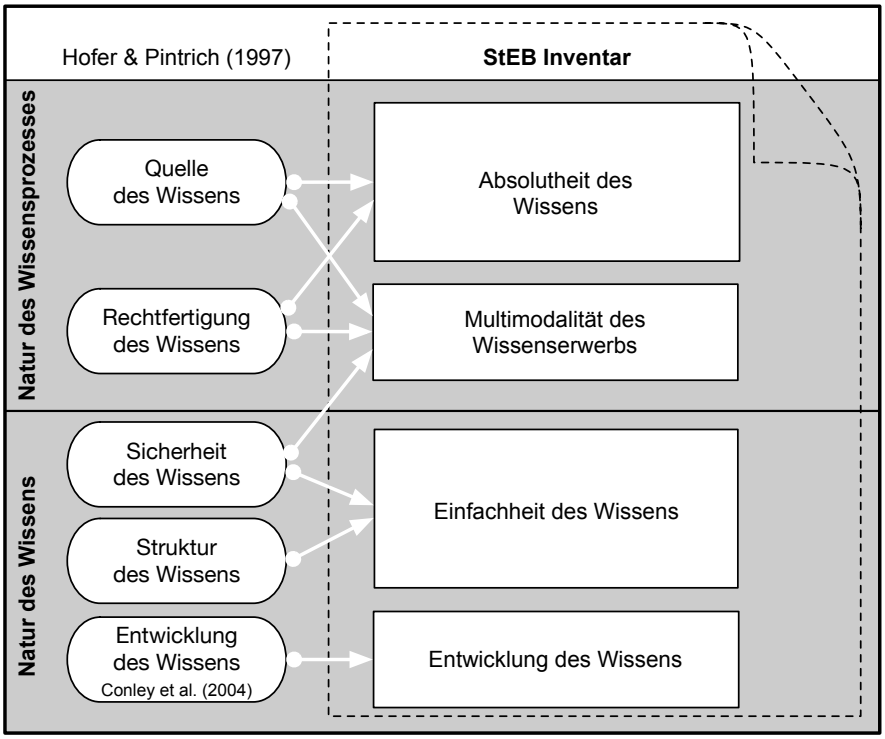


ABBILDUNG 6.4: Verortung der vier StEB-Dimensionen in den übergeordneten Bereichen epistemologischer Überzeugungen

Die Itemformulierung erfolgte zu den angenommenen Kernbereichen der epistemologischen Überzeugungen: *Quelle des Wissens*, *Sicherheit des Wissens*, *Struktur des Wissens* und *Rechtfertigung des Wissens*. Im Rahmen der ersten Studie ließen sich diese theoretisch angenommenen Kerndimensionen Hofer und Pintrichs (1997) im Testentwurf des StEB Inventars faktorenanalytisch jedoch nicht abbilden. Die *Entwicklung des Wissens*, die Conley et al. (2004) als von den Dimensionen der *Sicherheit* und *Komplexität des Wissens* unabhängige Dimension vorschlugen, konnte im Testentwurf des StEB Inventars faktorenanalytisch gezeigt werden. Betrachtet man allerdings die resultierenden Skalen im Kontext des zugrundegelegten Theoriegefüges (siehe Abb. 6.4), fällt auf, dass sie mit diesem vereinbar sind. Aus

Items zur *Quelle des Wissens* und *Rechtfertigung des Wissens* bildet sich die StEB-Dimension *Absolutheit des Wissens*, die im übergeordneten theoretischen Bereich *nature of knowing* anzusiedeln ist. Items der Skalen *Sicherheit des Wissens* und *Struktur des Wissens* bilden gemeinsam die StEB-Dimension *Einfachheit des Wissens*, die in den Bereich *nature of knowledge* fällt (siehe Abb. 6.4, S. 213). Die neue StEB-Dimension *Multimodalität des Wissenserwerbs* setzt sich aus Items zusammen, die ursprünglich den Dimensionen Quelle, Rechtfertigung und Sicherheit zugeordnet waren.

Eine sorgfältige Analyse der Iteminhalte der resultierenden Subskalen zeigt, dass sich alle vier Dimensionen, auch die ganz neu entstandene *Multimodalität des Wissenserwerbs*, im Hinblick auf das Konstrukt inhaltlich sinnvoll beschreiben lassen. Zusammengefasst lassen sich die vier resultierenden Dimensionen und ihre jeweiligen Pole folgendermaßen beschreiben:

**Überzeugungen zur Absolutheit des Wissens**

**Sophistizierter Pol:** Erziehungswissenschaftliche Studientexte stellen eine unzuverlässige Quelle für erziehungswissenschaftliches Wissen dar, deren Inhalte anzuzweifeln sind. Erziehungswissenschaftliches Wissen beim Lernen mit Studientexten ist relativ.

**Wenig sophistizierter Pol:** Erziehungswissenschaftliche Studientexte stellen eine verlässliche Quelle für erziehungswissenschaftliches Wissen dar, deren Inhalte man ohne zu hinterfragen glauben muss.

**Überzeugungen zur Multimodalität des Wissenserwerbs**

**Sophistizierter Pol:** Der eigene erziehungswissenschaftliche Wissenserwerb bedarf einer vielfältigen, elaborierten Auseinandersetzung mit erziehungswissenschaftlichen Studientexten. Der erziehungswissenschaftliche Wissenserwerb beim Lernen mit Studientexten ist demnach multimodal und betrifft die eigenständige Auseinandersetzung mit verschiedenen Wissensquellen, Forschungsmeinungen, Sichtweisen und Informationen aus verschiedenen Studientexten für das eigene Verstehen.

**Wenig sophistizierter Pol:** Für den erziehungswissenschaftlichen Wissenserwerb reicht eine einfache Auseinandersetzung mit den vorgegebenen erziehungswissenschaftlichen Studientexten aus. Das einfache Lesen eines Studientextes reicht als Auseinandersetzung aus; weitere Wissensquellen oder andere Forschungsmeinungen und Sichtweisen sind nicht relevant. Der erziehungswissenschaftliche Wissenserwerb aus Studientexten ist demnach unimodal.

**Überzeugungen zur Einfachheit des Wissens**

**Sophistizierter Pol:** Das Wissen in erziehungswissenschaftlichen Studientexten besteht aus unsicheren, untereinander verbundenen Konzepten. Erziehungswissenschaftliches Wissen aus Studientexten verfügt demnach über eine komplexe Struktur.

**Wenig sophistizierter Pol:** Das Wissen in erziehungswissenschaftlichen Studientexten besteht aus sicheren, diskreten, zu wissenden Fakten. Erziehungswissenschaftliches Wissen aus Studientexten verfügt demnach über eine einfache Struktur.

**Überzeugungen zur Entwicklung des Wissens**

**Sophistizierter Pol:** Das erziehungswissenschaftliche Wissen, das in Studientexten dargestellt wird, verändert sich über die Zeit, wodurch eine neue Auseinandersetzung mit den Inhalten nötig wird.

**Wenig sophistizierter Pol:** Das erziehungswissenschaftliche Wissen, das in Studientexten dargestellt wird, bleibt über die Zeit unverändert, weshalb eine erneute Auseinandersetzung mit den Inhalten nicht nötig wird.

Im Hinblick auf die Fragestellung lässt sich also schlussfolgern, dass die auf Basis der Studie 1 resultierenden Skalen tatsächlich mit dem Theoriegefüge vereinbar sind und inhaltlich die epistemologischen Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen aus Studientexten abdecken.

Im Sinne der Testentwicklung nach der klassischen Testtheorie (vgl. Kap. 4, ab S. 143) konnte in einem ersten Schritt anhand einer ersten Stichprobe mittels der EFA die Dimensionalität und Faktorstruktur des StEB Testentwurfes explorativ ermittelt werden. Im folgenden Kapitel wird nun beschrieben, wie diese Faktorstruktur der resultierenden 34-Item Version des StEB Inventars anhand einer neuen, größeren Stichprobe mit dem modellgenerierenden Ansatz der konfirmatorischen Faktorenanalyse geprüft wurde.

# 7 Konfirmatorische Modellbildung und Kreuzvalidierung (Studie 2)

Wie bereits in Kapitel 4 (ab S. 143) beschrieben, bietet es sich für die Testentwicklung nach der klassischen Testtheorie an, die EFA mit dem modellgenerierenden Ansatz der konfirmatorischen Faktorenanalyse zu kombinieren (Bühner, 2011). So können in einem ersten Schritt und anhand einer ersten Stichprobe (vgl. Kap. 6, ab S. 183) die Dimensionalität und Faktorstruktur des Testentwurfes explorativ ermittelt und in einem zweiten Schritt anhand einer neuen, größeren Stichprobe geprüft werden (vgl. Brown, 2015; Bühner, 2011; Fabrigar et al., 1999; R. B. Kline, 2016, S. 198).

In diesem Kapitel wird das auf Basis der ersten Stichprobe modifizierte StEB Inventar, im Rahmen einer zweiten Studie mit einer neuen, größeren Stichprobe geprüft (vgl. Brown, 2015; Bühner, 2011; Fabrigar et al., 1999; R. B. Kline, 2016, S. 198). Zum Einsatz kommt dazu der modellgenerierende Ansatz der konfirmatorischen Faktorenanalyse (CFA) (vgl. Jöreskog, 1993), da die CFA in dieser Art als ein Instrument zur Konstruktion und Evaluation von Skalen eingesetzt werden kann. Den Anfang bildet hier die Beschreibung der Durchführung und Datenerhebung. An dieser Stelle wird auch das Inventar zur Erfassung von Lernstrategien im Studium kurz vorgestellt (LIST-R, revidierte Version, Wild, 2011), da es neben dem StEB Inventar im Rahmen der Erhebung eingesetzt wurde. Dem schließen sich die Datenbereinigung und die deskriptive Analyse sowie die Analyse der fehlenden Werte an. Im Rahmen der folgenden faktorenanalytischen Auswertungen mittels CFAs wird geprüft, ob sich die Faktorstruktur replizieren lässt, die aus der ersten Studie resultiert. Im Anschluss wird eine Modelloptimierung vorgenommen, die über eine Teilung der Stichprobe in zwei Substichproben Schritt für Schritt kreuzvalidiert wird. In einem eigenen Teilkapitel werden im Folgenden die Item- und Skalenanalyse dargestellt. Das Kapitel schließt mit einer Diskussion der gesamten faktorenanalytischen Auswertung der zweiten Studie.

## 7.1 Durchführung

Die rekonstruierte 34-Item Version des StEB wurde Ende des Wintersemesters 2016/17 (von Mitte Januar bis Ende Februar) an einer zweiten, größeren Stichprobe von $N = 1051$ Lehramtsstudierenden verschiedener deutscher Universitäten eingesetzt. Der Fragebogen umfasste neben den StEB-Items zusätzlich einige demographische Angaben (persönlicher Code, Alter, Geschlecht, Hochschule, Studiengang, angestrebte Schulart, Studienbeginn, Fächerkombination, Fachsemester und Hochschulsemester). Zusätzlich wird außerdem das Inventar zur Erfassung von Lernstrategien im Studium in seiner revidierten Version (LIST-R,

Wild, 2011) eingesetzt. Das Ausfüllen des gesamten Fragebogens einschließlich des LIST-R nahm ca. 15-20 Minuten in Anspruch. Um eine für die Instrumententwicklung ausreichende Stichprobengröße zu gewährleisten, erfolgte die Erhebung sowohl vor Ort mittels einer Papierversion des Fragebogens als auch Online über LimeSurvey. Die Papierversion des Fragebogens wurde an den Universitäten Passau, Potsdam, Bremen sowie der Martin-Luther-Universität Halle-Wittenberg und der Pädagogischen Hochschule Freiburg eingesetzt. Die eingesetzte StEB-Fragebogenversion sowie der LIST-R können im Anhang A ab Seite 359 eingesehen werden. Wie schon bei der Ersterhebung erfolgte die Befragung in Seminaren und Vorlesungen aus den Studienbereichen der Schulpädagogik und pädagogischen Psychologie, da die Items sich domänenspezifisch auf Überzeugungen über Studientexte aus dem Bereich der Erziehungswissenschaft beziehen. Die Dozierenden, die die Fragebögen nach vorheriger Absprache per Post erhielten, teilten den Studierenden die Fragebögen in der jeweiligen Veranstaltung aus und gewährten die nötige Zeit zum Ausfüllen. In Passau besuchte die Autorin selbst 19 Veranstaltungen und teilte dort jeweils die Fragebögen aus. Mittels der Papierversion wurde ein Rücklauf von nahezu 100% erreicht. Insgesamt nahmen 704 Studierende an der Befragung mittels Papierversion teil.

Der Fragebogen wurde auf LimeSurvey als Onlineversion umgesetzt, um zusätzlich Studierende aus Veranstaltungen zu erreichen, bei denen der Einsatz der Papierversion aus organisatorischen Gründen nicht möglich gewesen wäre. Das betraf einerseits die Studierenden zweier Onlinevorlesungen („Einführung in die Medienpädagogik“ mit einer Teilnehmerzahl von 247 und „Bildungswissenschaft“ mit einer Teilnehmerzahl von 223) der Universität Passau sowie eine Vorlesung der PH Freiburg (Einführung in die Pädagogik und Didaktik der Primarstufe, angemeldetete Teilnehmerzahl 280) und mehrere Vorlesungen (aus den Bereichen Erziehungswissenschaft und pädagogischer Psychologie: Empirische Schul- und Bildungsforschung, Theorien der Sozialisationsforschung, Einführung in die Medienerziehung und Mediendidaktik, Sozialpsychologie I) der Universität Regensburg sowie der MLU-Halle und der Uni Bremen. Nach vorheriger Absprache mit den Dozierenden der betreffenden Veranstaltungen erhielten sie per Email einen Werbetext mit Link zum Fragebogen, der die Studierenden zur Teilnahme an der Befragung aufruft. Die Dozierenden wurden gebeten, den Werbetext in den jeweiligen die Veranstaltung begleitenden Onlinekurs (in StudIP, Moodle etc.) einzustellen und in der Veranstaltung um Teilnahme zu werden. Für die Werbung in der Veranstaltung wurde den Dozenten eine passende Powerpointfolie zur Verfügung gestellt. Circa zwölf Tage nach dem ersten Aufruf zur Teilnahme wurden die Studierenden in der Veranstaltung erneut zur Teilnahme aufgerufen sowie eine Erinnerungsnachricht im betreffenden Onlinekurs eingestellt. Ein entsprechender Erinnerungstext wurde den Dozierenden von der Autorin zur Verfügung gestellt. Die Zahl der Lehramtsstudierenden, die der Onlinefragebogen erreicht hat, ist schwer zu bestimmen, da die Zahl der Anmeldungen für Veranstaltungen sehr stark von der realen Teilnehmerzahl abweicht. Insgesamt erreichte der Onlinefragebogen schätzungsweise 1600 Studierende, von denen 347 den Fragebogen mindestens teilweise ausfüllten. Das entspricht einer Ausschöpfung von knapp 21,68%. Die Möglichkeit eines entstandenen Selection-Bias durch diese Online-Erhebung muss in Betracht gezogen werden.

## 7.2 Inventar zur Erfassung von Lernstrategien im Studium

Eine gekürzte, revidierte Version des LIST, der LIST-R (Wild, 2011), wurde im Rahmen dieser Arbeit zur Erfassung des durch die Lehramtsstudierenden berichteten Lernstrategieeinsatzes im Bereich ihrer erziehungswissenschaftlichen Studienanteile eingesetzt. Beim LIST handelt es sich um eine speziell für den Hochschulbereich entwickelte, deutsche Adaption des MSLQ, dessen Konzeption bereits in Kapitel 2.8.2 (in Abb. 2.10, S. 98) dargestellt wurde.

Der LIST-R wurde dem LIST, der über 77 Items verfügt, aus Gründen der Erhebungsökonomie vorgezogen. Er deckt mit seinen 53 Items die drei übergeordneten Bereiche *kognitive Lernstrategien*, *metakognitive Lernstrategien* und *ressourcenbezogene Strategien* ab. In Tabelle 7.1 wird der Aufbau des LIST-R im Überblick dargestellt. Der Bereich der kognitiven Lernstrategien umfasst die drei Skalen *Organisieren*, *Wiederholen* und *Elaborieren* mit je sieben Items zu diesen drei Strategiearten. Die metakognitiven Lernstrategien werden über eine Skala mit insgesamt zehn Items erfasst. Drei dieser Items beziehen sich auf die aktive Planung und sieben auf die Regulation des eigenen Lernverhaltens. Der Bereich der ressourcenbezogenen Strategien besteht aus fünf Subskalen mit je vier Items zu *Lernzeit*, *Literatur*, *Online-Ressourcen* und *kooperativem Lernen* sowie sechs Items zur *Lernumgebung* (siehe Tab. 7.1). Die Items des LIST-R sind im Anhang C in den Tabellen C.1 und C.2 (ab S. 390) einzusehen.

TABELLE 7.1: Aufbau des LIST-R mit Item-Skalen-Zuordnung, Itemanzahl je Skala und Reliabilität $\alpha$

| Lernstrategiebereiche/ Lernstrategieskalen | Label nach Wild (2011) | Item-anzahl | $\alpha$ |
|---|---|---|---|
| kognitiv | | | |
| Organisieren | Inhalte organisieren | 7 | .79 |
| Wiederholen | Inhalte wiederholen und einprägen | 7 | .74 |
| Elaborieren | Inhalte verbinden und prüfen | 7 | .82 |
| metakognitiv | | | |
| Metakognitive Lernstrategien (Subsets: Meine Planung, Meine Qualitätskontrolle) | Lernen organisieren | 3/7 | .81 |
| ressourcenbezogen | | | |
| Zeitmanagement | Lernzeit | 4 | .86 |
| Lernumgebung | Lernumgebung | 6 | .79 |
| Literatur | Literatur nutzen | 4 | .86 |
| Online-Ressourcen | Online-Quellen nutzen | 4 | .78 |
| Kooperatives Lernen | gemeinsam lernen | 4 | .86 |

Die folgende Beschreibung der Skalen stammt aus dem Kurzmanual zum LIST-R (Wild, 2011, S. 7):

**Wiederholungsstrategien** Studientätigkeiten, die auf das Einprägen von Fakten und Regeln durch Wiederholung bezogen sind. Sie dienen in erster Linie der Vorbereitung auf Prüfungen.

**Organisationsstrategien:** Studientätigkeiten, die durchgeführt werden, um die zu lernenden Inhalte in geeigneter Weise zu strukturieren und zu organisieren. Dies umfasst u.a. das Erstellen von Zusammenfassungen und Gliederungen oder das Anfertigen von Tabellen und Schaubildern.

**Elaborationsstrategien:** Studientätigkeiten, die auf ein sinnhaftes und reflektiertes Verstehen des Stoffes ausgerichtet sind. Dies geschieht, indem neue Inhalte systematisch auf das bereits vorhandene Wissen bezogen werden. Hierzu zählen auch das kritische Durchdenken von Aussagen und Argumentationen.

**Metakognitive Strategien:** Die Skala erfasst metakognitive Strategien zur aktiven Planung und Regulation des eigenen Lernverhaltens. Dies bezieht sich u.a. auf die Planung konkreter Arbeitsschritte, die für eine bevor stehende Lernphase notwendig sind. Weiterhin erfragt diese Skala Tätigkeiten, die zur Überwachung von Lernproblemen dienen. Ein dritter Aspekt bezieht sich auf die flexible Veränderung des Lernverhaltens bei Verständnisproblemen.

**Lernzeit:** Die Skala erfasst, inwieweit eine explizite Zeitplanung angestrebt wird.

**Lernumgebung:** Die Skala erfasst, inwieweit eine lernförderliche Lernumgebung aktiv geschaffen oder gesucht wird.

**Literatur:** Die Skala erfasst, inwieweit bei Bedarf auf zusätzliche Literatur zurückgegriffen wird.

**Onlineressourcen:** Die Skala erfasst, inwieweit bei Bedarf auf Angebote des Internets zurückgegriffen wird (neue Skala).

**Kooperatives Lernen:** Die Skala erfasst das Ausmaß kooperativen Lernens. Sie umfasst zum einen verschiedene Formen gemeinsamer Arbeit, sie umfasst aber auch Formen einseitiger Inanspruchnahme von Studienkolleg(inn)en (ehemals genannt: Lernen mit Studienkollegen).

## 7.3 Datenbereinigung

Die Datensätze der insgesamt 1051 Befragten (704 Papierversion, 347 Onlinebefragung) wurden zunächst gesichtet. Über Limesurvey gingen insgesamt 347 Datensätze ein, die Eintragungen enthielten. Bei 63 davon war die Bearbeitung jedoch direkt nach oder während der demografischen Angaben abgebrochen worden, sodass keinerlei Angaben zu den StEB- oder LIST-Items vorliegen und die Daten somit nicht verwertbar sind. Bei weiteren 42 Fragebögen wurden die demografischen Angaben zwar vollständig ausgefüllt, die Bearbeitung des Fragebogens jedoch während der StEB-Angaben abgebrochen. Auch diese Bögen sind nicht verwertbar, da sie über sehr unvollständige StEB-Angaben verfügen. Die Daten einer Person mussten wegen einem deutlich erkennbaren Ankreuzmuster ausgeschlossen werden. Diese Person hat ab StEB-Item Nummer sieben über den gesamten Fragebogen hinweg nur noch die drei angekreuzt. Vier weitere Datensätze konnten nicht verwendet werden, da dieselben Personen, wie sich mittels des persönlichen Codes feststellen ließ, bereits an der ersten Erhebung teilgenommen hatten. Demnach verbleiben von den 347 online ausgefüllten Fragebögen 237 verwertbare Datensätze.

704 Personen wurden über die Papierversion des Fragebogens direkt in einer Veranstaltung befragt. Die Daten von 16 Personen wurden aus dem Datensatz entfernt, da diese Personen, wie anhand des Codes identifiziert werden konnte, bereits an der ersten Erhebung teilgenommen hatten. Ein Datensatz wurde wegen einer hohen Anzahl fehlender Werte entfernt. Die betreffende Person hatte 17 der 34 StEB-Items nicht bearbeitet. Ein weiterer Datensatz wurde entfernt, da die befragte Person keinerlei demografische Angaben gemacht hat. 686 verwertbare Datensätze verbleiben von der Befragung in der Papierversion.

## 7.4 Deskriptive Datenanalyse

Nach der Bereinigung der Daten beträgt die Stichprobe $N$ = 923 (344 männlich, 574 weiblich, 5 keine Angabe) Lehramtsstudierende im durchschnittlichen Alter von 22.2 Jahren ($SD$ = 3.88). Im Zusammenhang mit der Prüfung der Voraussetzungen der konfirmatorischen Faktorenanalyse (siehe Kap. 7.5.1, ab S. 229) müssen weitere acht Datensätze aus der Stichprobe ausgeschlossen werden. Dadurch beträgt die Stichprobe für die Analysen schlussendlich $N$ = 915. Der Übersichtlichkeit halber, wird an dieser Stelle der abschließende Datensatz mit $N$ = 915 deskriptiv dargestellt, da dieser für die weiteren Analysen relevant ist.

Die Stichprobe besteht aus $N$ = 915 (341 männlich, 569 weiblich, 5 keine Angabe) Lehramtsstudierenden mit einem durchschnittlichen Alter von 22.0 Jahren ($SD$ = 3.34). Davon wurden 685 direkt in einer Veranstaltung mittels der Papierversion des Fragebogens befragt, während 230 den Fragebogen über LimeSurvey online ausfüllten. Rund 84% der Befragten gaben in Bezug auf ihren angestrebten Abschluss an, klassisches Lehramt zu studieren, während die übrigen 145 zusätzlich einen Bachelor- oder Masterabschluss erwerben.

Es wurden Lehramtsstudierende sechs verschiedener deutscher Hochschulen in den fünf Bundesländern Bayern, Baden-Württemberg, Sachsen-Anhalt, Brandenburg und Bremen befragt. Mit rund 45% studiert der größte Anteil der Stichprobe an der Universität Passau. Nur 14 Studierende konnten an der Universität Potsdam erreicht werden (siehe Tab. 7.2). Von denen in Lehrveranstaltung an der PH Freiburg befragen Studierenden, gaben elf Personen an, an der Universität Freiburg und nicht der PH Freiburg zu studieren.

TABELLE 7.2: Häufigkeiten und prozentuale Anteile nach Hochschule (nach Ausschluss der Ausreißer)

| Hochschule | Probanden | % |
|---|---|---|
| Universität Passau | 404 | 44,7 |
| MLU Halle-Wittenberg | 174 | 19,3 |
| PH Freiburg | 188 | 20,8 |
| Universität Regensburg | 61 | 6,8 |
| Universität Bremen | 51 | 5,6 |
| Universität Potsdam | 14 | 1,6 |
| Universität Freiburg | 11 | 1,2 |
| Gesamt | 903 | 100 |

Der Großteil der Studierenden, knapp 40%, studiert Lehramt für Grundschulen. Die zweithäufigste Schulart ist Gymnasium, gefolgt von der Sekundarschule und Mittelschule (siehe Tab. 7.3).

TABELLE 7.3: Häufigkeiten und prozentuale Anteile nach studierter Schulart (nach Ausschluss der Ausreißer)

| Schulart | Probanden | % |
|---|---|---|
| Grundschule | 352 | 38,6 |
| Hauptschule | 3 | 0,3 |
| Realschule | 100 | 11 |
| Mittelschule | 101 | 11,1 |
| Gymnasium | 194 | 21,3 |
| Förderschule | 17 | 1,9 |
| Sekundarschule | 130 | 14,3 |
| Grundschule & Förderschule | 15 | 1,6 |
| Gesamt | 912 | 100 |

Die Befragten studieren im Durchschnitt im dritten Fachsemester ($M$ = 3.12, $SD$ = 2.72). Seit durchschnittlich 4.57 Semestern ($SD$ = 3.90) sind sie bereits ingesamt an einer Hochschule eingeschrieben. Aus Tabelle 7.4 auf Seite 223 geht hervor, dass rund 50% der Befragten ihr Lehramtsstudium zum Wintersemester 2016/17 begonnen haben. Zum Zeitpunkt der Befragung – Januar bis Februar 2017 – befanden sie sich somit am Ende ihres ersten Fachsemesters. Die andere Hälfte der Befragten befindet sich in höheren Fachsemestern, mit 134 im dritten und 109 im fünften Semester.

TABELLE 7.4: Häufigkeiten und prozentuale Anteile nach Studienbeginn und Fachsemester (nach Ausschluss der Ausreißer)

| Studienbeginn | Fachsemester | Probanden | % |
|---|---|---|---|
| WiSe 10/11 | 13 | 4 | 0,4 |
| SoSe 11 | 12 | 2 | 0,2 |
| WiSe 11/12 | 11 | 12 | 1,3 |
| SoSe 12 | 10 | 5 | 0,5 |
| WiSe 12/13 | 9 | 29 | 3,2 |
| SoSe 13 | 8 | 5 | 0,5 |
| WiSe 13/14 | 7 | 89 | 9,8 |
| SoSe 14 | 6 | 22 | 2,4 |
| WiSe 14/15 | 5 | 109 | 12 |
| SoSe 15 | 4 | 17 | 1,9 |
| WiSe 15/16 | 3 | 134 | 14,7 |
| SoSe 16 | 2 | 29 | 3,2 |
| WiSe 16/17 | 1 | 455 | 49,9 |
| Gesamt | | 912 | 100 |

Rund 35% (303) der Befragten, die dazu eine Auskunft gaben (44 k.A. von $N$ = 915), sind „echte“ Studienanfänger im ersten Hochschulsemester. 17,1% (149) befinden sich im zweiten oder dritten Hochschulsemester, 15,4% (134) im vierten oder fünften, 16,2% (141) im sechsten oder siebten Hochschulsemester. Die übrigen 144 Personen (16,5%) sind bereits seit über sieben Semestern an einer Hochschule eingeschrieben.

Die Häufigkeitsverteilungen der in der Stichprobe vertretenen Studienfächer und dementsprechend späteren Unterrichtsfächer der Lehramtsstudierenden gehen aus den Tabellen 7.5 und 7.6 auf Seite 224 hervor. Die fünf am häufigsten vertretenen Erstfächer sind Deutsch, Mathematik, Englisch sowie Geschichte und Geographie. Die Fächer Mathematik und Deutsch sind auch als Zweit- und Drittfach mit Abstand am häufigsten vertreten. Kunst stellt das häufigste Viertfach dar. Zu berücksichtigen ist bei der hier berichteten Gesamtzahl an angegebenen Fächern, dass die Zahl der insgesamt zu wählenden Unterrichtfächer je nach Schulart verschieden ist. Während Studierende des Lehramts für Grundschulen sehr häufig vier Fächer studieren, sind beim Lehramt für Gymnasium zwei Fächer die Regel.

TABELLE 7.5: Häufigkeiten und prozentuale Anteile der studierten Erst- und Zweitfächer (nach Ausschluss der Ausreißer)

| Erstfach | | | Zweitfach | | |
|---|---|---|---|---|---|
| Rangfolge der Fächer | Probanden | % | Rangfolge der Fächer | Probanden | % |
| Deutsch | 230 | 25,5 | Mathematik | 193 | 21,5 |
| Mathematik | 134 | 14,8 | Deutsch | 187 | 20,8 |
| Englisch | 142 | 15,6 | Geschichte | 83 | 9,3 |
| Geschichte | 63 | 6,9 | Englisch | 53 | 5,9 |
| Geographie | 62 | 6,8 | Sport | 45 | 5 |
| Sport | 36 | 4 | Geographie | 37 | 4,1 |
| Religion | 54 | 6 | Biologie | 26 | 2,9 |
| Biologie | 29 | 3,2 | AuG* | 25 | 2,8 |
| Kunst | 18 | 2 | Kunst | 24 | 2,7 |
| Musik | 17 | 1,9 | Wirtschaft | 23 | 2,6 |
| Sozialkunde | 16 | 1,8 | Französisch | 19 | 2,1 |
| Französisch | 14 | 1,5 | Chemie | 17 | 1,9 |
| Wirtschaft | 14 | 1,5 | Musik | 16 | 1,8 |
| Inklusive Pädagogik | 11 | 1,2 | Sozialkunde | 15 | 1,7 |
| Latein | 10 | 1,1 | Physik | 14 | 1,6 |
| AuG* | 6 | 0,7 | Ethik | 11 | 1,2 |
| Informatik | 6 | 0,7 | Sachunterricht | 11 | 1,2 |
| Sonstige | 58 | 6,4 | Sonstige | 98 | 10,9 |
| Gesamt | 908 | 100 | | 897 | 100 |

* Alltagskultur und Gesundheit

TABELLE 7.6: Häufigkeiten und prozentuale Anteile der studierten Dritt- und Viertfächer (nach Ausschluss der Ausreißer)

| Drittfach | | | Viertfach | | |
|---|---|---|---|---|---|
| Rangfolge der Fächer | Probanden | % | Rangfolge der Fächer | Probanden | % |
| Mathematik | 114 | 22 | Kunst | 98 | 28,7 |
| Deutsch | 74 | 14,3 | Sport | 84 | 24,6 |
| Englisch | 39 | 7,5 | Musik | 48 | 14,1 |
| Arbeitslehre | 32 | 6,2 | Religion | 14 | 4,1 |
| Sozialkunde | 24 | 4,6 | Deutsch | 12 | 3,5 |
| Sport | 23 | 4,4 | Mathematik | 11 | 3,2 |
| Sachunterricht | 19 | 3,7 | Englisch | 10 | 2,9 |
| Kunst | 18 | 3,5 | Ezw* | 8 | 2,3 |
| Geschichte | 17 | 3,3 | Sachunterricht | 8 | 2,3 |
| Geographie | 16 | 3,1 | Geographie | 7 | 2,1 |
| Musik | 14 | 2,7 | Geschichte | 7 | 2,1 |
| Ethik | 14 | 2,7 | Arbeitslehre | 7 | 2,1 |
| Sonstige | 115 | 22,2 | Sonstige | 27 | 7,9 |
| Gesamt | 519 | | | 341 | 100 |

* Erziehungswissenschaft

Tabelle 7.7 zeigt wie häufig jedes der Fächer unter den $N = 915$ Lehramtsstudierenden insgesamt vorkommt. Hier wird nicht nach Erst- bis Viertfach unterschieden. Die Fächer werden getrennt für die drei Bereiche der Geistes-, Natur- und Sozialwissenschaft aufgeführt. Die Fächer, die in dieser Stichprobe insgesamt am häufigsten studiert werden sind Deutsch, gefolgt von Mathematik, Englisch, Sport, Geschichte, Kunst und Geographie. Neben Englisch sind andere Sprachen wie Spanisch und Italienisch kaum vertreten. Auch beispielsweise die Fächer Informatik und Technik werden nur von wenigen Lehramtsstudierenden dieser Stichprobe studiert.

TABELLE 7.7: Häufigkeiten der Fächer und Zuordnung der Fächer zu Bereichen ($N = 915$)

| Geistes-wissenschaft | Häufigkeit | Natur-wissenschaft | Häufigkeit | Sozial-wissenschaft | Häufigkeit |
|---|---|---|---|---|---|
| Deutsch | 513 | Mathematik | 463 | Sport | 190 |
| Englisch | 244 | Geographie | 123 | Sozialkunde | 61 |
| Geschichte | 170 | Biologie | 63 | Arbeitslehre | 44 |
| Kunst | 165 | Sachunterricht | 44 | SPF | 44 |
| Religion | 108 | Chemie | 22 | AKG | 38 |
| Musik | 95 | Physik | 21 | EZW | 15 |
| Ethik | 29 | Technik | 14 | Politik | 11 |
| Französich | 34 | AWT | 13 | Psychologie | 2 |
| Latein | 14 | Astronomie | 8 | Soziologie | 1 |
| Philosophie | 10 | Informatik | 9 | | |
| Spanisch | 8 | Erdkunde | 6 | | |
| Italienisch | 4 | | | | |
| Russisch | 1 | | | | |

AWT Arbeit Wirtschaft Technik, AKG Alltagskultur und Gesundheit, EZW Erziehungswissenschaft, SPF sonderpädagogische Fächer

### 7.4.1 Deskriptive Analyse der StEB-Items

Die folgende Tabelle 7.8 auf Seite 226 stellt die StEB-Überzeugungsitems deskriptiv dar. Die Itemmittelwerte reichen von minimal 2.00 für Item sw_15_u bis zu einem Mittelwert von maximal 4.25 für Item gw_17. Item gw_17 verfügt gleichzeitig mit $SD = 0.77$ über die geringste Standardabweichung, Item gw_13_u mit $SD = 1.14$ dagegen über die höchste. Bei allen Items bis auf einem wird die volle Breite des Antwortformats von 1 (trifft gar nicht zu) bis 5 (trifft voll zu) genutzt. Keiner der 915 Befragten lehnte die Aussage von Item gw_17 („Ich bin offen gegenüber Ansichten in pädagogischen Studientexten, die ich bisher noch nicht kannte.“) voll ab, das heißt kreuzte 1 („trifft gar nicht zu“) an.

TABELLE 7.8: Deskriptive Statistik der StEB-Items (nach Ausschluss der Ausreißer)

| Items | *n* | *M* | *SD* | *Md* | Min | Max | Range | skew | kurtosis | *SE* |
|---|---|---|---|---|---|---|---|---|---|---|
| gw_17 | 914 | 4.25 | 0.77 | 4 | 2 | 5 | 3 | -0.85 | 0.37 | .03 |
| rw_33_u | 915 | 2.80 | 0.91 | 3 | 1 | 5 | 4 | 0.21 | -0.15 | .03 |
| gw_3_u | 915 | 2.37 | 1.03 | 2 | 1 | 5 | 4 | 0.45 | -0.46 | .03 |
| ew_5_u | 914 | 2.87 | 0.92 | 3 | 1 | 5 | 4 | 0.13 | -0.34 | .03 |
| rw_11 | 915 | 3.96 | 0.94 | 4 | 1 | 5 | 4 | -0.79 | 0.19 | .03 |
| sw_2_u | 913 | 2.06 | 0.94 | 2 | 1 | 5 | 4 | 0.63 | -0.15 | .03 |
| qw_21_u | 914 | 2.94 | 0.81 | 3 | 1 | 5 | 4 | 0.36 | 0.48 | .03 |
| ew_13 | 914 | 3.88 | 0.81 | 4 | 1 | 5 | 4 | -0.39 | -0.18 | .03 |
| sw_11_u | 915 | 2.38 | 0.95 | 2 | 1 | 5 | 4 | 0.28 | -0.43 | .03 |
| rw_18_u | 915 | 2.86 | 0.98 | 3 | 1 | 5 | 4 | 0.32 | -0.43 | .03 |
| rw_27 | 915 | 4.04 | 0.88 | 4 | 1 | 5 | 4 | -0.72 | 0.04 | .03 |
| ew_8 | 913 | 3.22 | 0.86 | 3 | 1 | 5 | 4 | 0.02 | -0.43 | .03 |
| rw_8_u | 915 | 2.32 | 0.79 | 2 | 1 | 5 | 4 | 0.40 | 0.19 | .03 |
| sw_17_u | 913 | 2.26 | 0.97 | 2 | 1 | 5 | 4 | 0.58 | -0.14 | .03 |
| sw_22 | 915 | 3.95 | 0.79 | 4 | 1 | 5 | 4 | -0.60 | 0.35 | .03 |
| sw_15_u | 912 | 2.00 | 0.97 | 2 | 1 | 5 | 4 | 0.60 | -0.47 | .03 |
| rw_12 | 915 | 3.64 | 0.89 | 4 | 1 | 5 | 4 | -0.37 | -0.22 | .03 |
| ew_3 | 914 | 3.40 | 0.81 | 3 | 1 | 5 | 4 | 0.02 | -0.29 | .03 |
| qw_25_u | 915 | 3.38 | 0.96 | 3 | 1 | 5 | 4 | -0.07 | -0.41 | .03 |
| rw_36_u | 912 | 2.98 | 1.13 | 3 | 1 | 5 | 4 | 0.03 | -0.73 | .04 |
| ew_10 | 912 | 3.54 | 0.86 | 4 | 1 | 5 | 4 | -0.22 | -0.43 | .03 |
| gw_9 | 914 | 3.84 | 0.95 | 4 | 1 | 5 | 4 | -0.59 | -0.23 | .03 |
| qw_26_u | 915 | 3.04 | 1.09 | 3 | 1 | 5 | 4 | 0.06 | -0.77 | .04 |
| ew_12 | 914 | 3.73 | 0.85 | 4 | 1 | 5 | 4 | -0.15 | -0.41 | .03 |
| qw_11 | 914 | 3.85 | 0.94 | 4 | 1 | 5 | 4 | -0.64 | -0.04 | .03 |
| sw_3_u | 913 | 3.83 | 0.92 | 4 | 1 | 5 | 4 | -0.78 | 0.56 | .03 |
| qw_13_u | 915 | 3.27 | 0.95 | 3 | 1 | 5 | 4 | -0.11 | -0.40 | .03 |
| sw_9 | 913 | 3.77 | 0.82 | 4 | 1 | 5 | 4 | -0.49 | 0.19 | .03 |
| gw_13_u | 914 | 2.87 | 1.14 | 3 | 1 | 5 | 4 | 0.02 | -0.80 | .04 |
| ew_2 | 914 | 3.54 | 0.96 | 4 | 1 | 5 | 4 | -0.33 | -0.42 | .03 |
| gw_4_u | 913 | 2.46 | 0.84 | 2 | 1 | 5 | 4 | 0.33 | 0.00 | .03 |
| gw_16 | 914 | 3.31 | 0.91 | 3 | 1 | 5 | 4 | -0.17 | -0.32 | .03 |
| rw_29_u | 914 | 2.92 | 1.01 | 3 | 1 | 5 | 4 | 0.25 | -0.52 | .03 |
| ew_16_u | 914 | 3.57 | 0.82 | 3 | 1 | 5 | 4 | 0.28 | -0.51 | .03 |

### 7.4.2 Analyse fehlender Werte

Die Variablen Alter, Geschlecht, Universität, Abschluss, Schulart, Studienbeginn, Hochschulsemester, Fachsemester und Erstfach bis Viertfach wurden gemeinsam mit den 34 Items zu den epistemologischen Überzeugungen mittels des Little Tests (Little & Rubin, 1987) auf eine zugrundeliegende Systematik der fehlenden Werte geprüft (*missing-completely-at-random*, MCAR) (R Paket: BaylorEdPsych). Mit $\chi^2$= 2033.762, $df$ = 1884, $p$ = .008 kann davon ausgegangen werden, dass die MCAR-Bedingung nicht erfüllt ist (Little & Rubin, 2002). Für die Auswertung relevant sind jedoch die fehlenden Werte bei den StEB-Items, da diese Items mittels der konfirmatorischen Faktorenanalyse analysiert werden sollen. Aus der Spalte *n* in Tabelle 7.8 auf Seite 226 geht hervor, wieviele Personen das jeweilige Item beantwortet haben. Hier zeigt sich, dass je Item maximal drei Werte fehlen. Je drei Personen haben Item sw_15_u, rw_36_u und ew_10 nicht beantwortet. Für alle anderen Items fehlen weniger als drei Werte. Berechnet man den Little Test ausschließlich über die 34 StEB-Items, ergibt sich ebenfalls ein signifikantes Ergebnis, das heißt die MCAR-Bedingung ist verletzt ($\chi^2$ = 678.6658, $df$ = 614, $p$ = .036).

Das Ergebnis des Little Tests ist an dieser Stelle allerdings mit Vorsicht zu betrachten, da zwei Einschränkungen für seine Verlässlichkeit vorliegen. Einerseits ist er abhängig von der Stichprobengröße und lehnt somit die MCAR Annahme bei großen Stichproben leichter ab, und zweites setzt er eine Normalverteilung der Daten voraus (Enders, 2010), die in diesem Fall nicht gegegeben ist, wie im Zusammenhang mit den Voraussetzungen der CFA (ab S. 229) erläutert wird.

Die nicht direkt zu testende MAR-Bedingung lässt sich nur indirekt einschätzen (Enders, 2010). Auffällig ist, dass der prozentuale Anteil fehlender Werte für die StEB-Items mit maximal 0,3% sehr gering ist. Dies und die Tatsache, dass sich die fehlenden Werte nicht bei bestimmten Items häufen, spricht für das Vorliegen der MAR-Bedingung. Wie bereits in Kapitel 6.2.2 (ab S. 191) beschrieben, gibt es drei übergeordnete Arten des Umgangs mit fehlenden Werten: Ausschluss (*deletion*), einfache Imputation (*single imputation*) und multiple Imputation (Enders, 2010). Darüber hinaus stellt Maximum-Likelihood Ansatz (auch genannt: *full information maximum likelihood*, FIML) ein robustes Verfahren dar.

Neben der multiplen Imputation wird FIML als state-of-the-art Methode zum Umgang mit fehlenden Werten beschrieben (Schafer & Graham, 2002), da es bei Vorliegen von MCAR oder MAR (nicht jedoch bei MNAR) zu verlässlichen Schätzungen führt (Enders, 2010, S. 13). Zum Umgang mit den fehlenden Werten in den StEB-Daten wurde die full-information-Maximum-Likelihood-Methode (FIML) gewählt. Es handelt sich um ein robustes Verfahren zum Umgang mit fehlenden Werten, das bei zugrundeliegendem MAR Mechanismus und auch bei MCAR zu genauen Parameterschätzungen führt (Enders, 2010, S. 125) und ist damit traditionellen Verfahren wie der einfachen Imputation deutlich überlegen.

Wenn fehlende Daten vorliegen, muss die Berechnung der Standardfehler mittels ML angepasst werden. Es muss unterschieden werden zwischen Standardfehlern, die auf der beobachteten Datenmatrix basieren und solchen, die auf der geschätzten (*expected*) Datenmatrix basieren. Aus der beobachteten Datenmatrix resultieren Standardfehler, die bei MAR-Daten angemessen sind, während die geschätzte Datenmatrix die MCAR-Annahme voraussetzt (Enders, 2010, S. 125). In der folgenden faktorenanalytischen Auswertung wurde dementsprechend die beobachtete Datenmatrix zur Ermittlung der Standardfehler verwendet.

## 7.5 Faktorenanalytische Auswertung der Validierungsstichprobe

Die Daten der ersten Stichprobe wurden mittels der exploratorischen Faktorenanalyse ausgewertet, um die Faktorstruktur des zukünftigen Fragebogens zu untersuchen und die Itemauswahl zu optimieren (Brown, 2015, S. 10). Im darauf folgenden Schritt der Testentwicklung wird die konfirmatorische Faktorenanalyse (*confirmatory factor analysis*, CFA) eingesetzt. Sie „dient dazu, theoretisch oder empirisch gut fundierte Modelle auf ihre empirische Passung mit den Daten hin zu testen“ Bühner (2011, S. 380). Mittels der CFA wird die latente Struktur des Testinstruments (Brown, 2015, S. 1) hypothesengeleitet anhand einer neuen, größeren Stichprope (Bühner, 2011) erkundet. In diesem Zusammenhang werden durch die CFA vordefinierte Annahmen über die Anzahl der zugrundeliegenden Dimensionen des Fragebogens (Faktoren) sowie die Zuordnung der Items zu den Faktoren (Faktorladungen) überprüft (Brown, 2015, S. 1). Diese vordefinierten Annahmen basieren auf der der Instrumententwicklung zugrundeliegenden Theorie und den spezifischen Ergebnissen aus dem vorherigen Entwicklungsschritt, d.h. den Ergebnissen der EFA. Die CFA dient auch dazu, festzulegen, wie im fertigen Erhebungsinstrument die individuellen Testwerte von Versuchspersonen ermittelt werden sollten. Wird eine einfaktorielle Struktur (Eindimensionalität) ermittelt, kann ein Gesamttestwert über alle Items hinweg berechnet werden. Wird jedoch eine multifaktorielle Struktur identifiziert (mehr als zwei Dimensionen), legt das Muster der Faktorladungen der CFA fest, wie der Testwert auf Basis von Subskalen des Instruments ermittelt wird. Die Anzahl der identifizierten Faktoren entspricht der Anzahl der Subskalen, die das Testinstrument enthält. Das Muster der Item-Faktor-Beziehungen legt dabei fest, welches Item welcher Subskala angehört. Dementsprechend muss dann je Subskala ein eigener Testwert über die entsprechend zugehörigen Items berechnet werden (Brown, 2015, S. 1 f.). Die CFA eignet sich im Zusammenhang mit der Instrumententwicklung auch für die Analyse anderer psychometrischer Eigenschaften wie der Schätzung der Skalenreliabilität. Die CFA wird auch im Zusammenhang mit der Validierung von Testwertinterpretationen bzw. nach dem früheren Validitätsverständnis zur Konstruktvalidierung eingesetzt (Brown, 2015, S. 2).

Da der genauere Ablauf der CFA und die einzelnen Entscheidungsschritte bereits in Kapitel 4.3.2 (ab S. 160) beschrieben wurde, liegt der Fokus im Folgenden auf den in dieser Auswertung konkret getroffenen Entscheidungen und vorgenommenen Maßnahmen.

### 7.5.1 Prüfung der Voraussetzung für die CFA

Bevor die konfirmatorische Faktorenanalyse durchgeführt wird, sind die Voraussetzungen bezüglich Ausreißer, des Skalenniveaus, der Verteilung, der Kollinearität, der Stichprobengröße und der benötigten Anzahl von Indikatoren pro latenter Variable zu berücksichtigen (Bühner, 2011).

Vor der Durchführung der CFA müssen Ausreißer beispielsweise über die Inspektion von Histogrammen oder Streudiagrammen identifiziert und ausgeschlossen werden, da diese die Zusammenhänge zwischen den Variablen verzerren und verhindern, dass sich die

Ergebnisse der CFA sinnvoll interpretieren lassen (Bühner, 2011, S. 432). Diese Voraussetzung der CFA wird zuerst geprüft, da der zu analysierende Datensatz ($N$ = 923) sich durch den Ausschluss von Ausreißern unter Umständen noch ändert, was dazu führen würde, dass nachfolgende Tests, die im Rahmen der Prüfung der Voraussetzung nötig sind, nach dem Ausschluss der Ausreißer wiederholt werden müssten. Da in einer Testentwicklung angestrebt wird, dass die Stichprobe repräsentativ für die Zielgruppe des Tests ist, wird der Datensatz ($N$ = 923) bezüglich Ausreißern in Alter und Studiensemester untersucht. Für die Erstellung der Streudiagramme wurden zunächst die individuellen Skalenmittelwerte der vier angenommenen Skalen, die aus der ersten Stichprobe resultierten, berechnet. Für die Erstellung der Streudiagramme wurden die umgekehrt formulierten Items zunächst umgepolt, da die Skala *Entwicklung des Wissens* Items beider Formulierungsrichtungen enthält. Für alle Items entsprach somit nach der Umpolung eine hohe Zustimmung sophistizierten epistemologischen Überzeugungen. Nach Abschluss dieser Prüfung auf Ausreißerwerte wurde wieder ein Datensatz ohne umkodierte StEB-Items verwendet. Anhand des Streudiagramms 7.1 zur Subkala *Absolutheit des Wissens* mit dem Alter ist zu erkennen, dass die beiden Fälle mit einem Alter über 42 Jahren Ausreißer darstellen. Auch Personen mit mehr als 15 Fachsemestern (siehe Abb. 7.2, S. 231) und solche mit mehr als 25 Hochschulsemestern (siehe Abb. 7.3, S. 231) wurden als Ausreißer identifiziert. Die Daten der Personen, die kein Alter angaben, konnten beibehalten werden, da sie sich bezüglich ihrer vorhandenen Angaben zu Fachsemester und Hochschulsemester im Normbereich befanden.

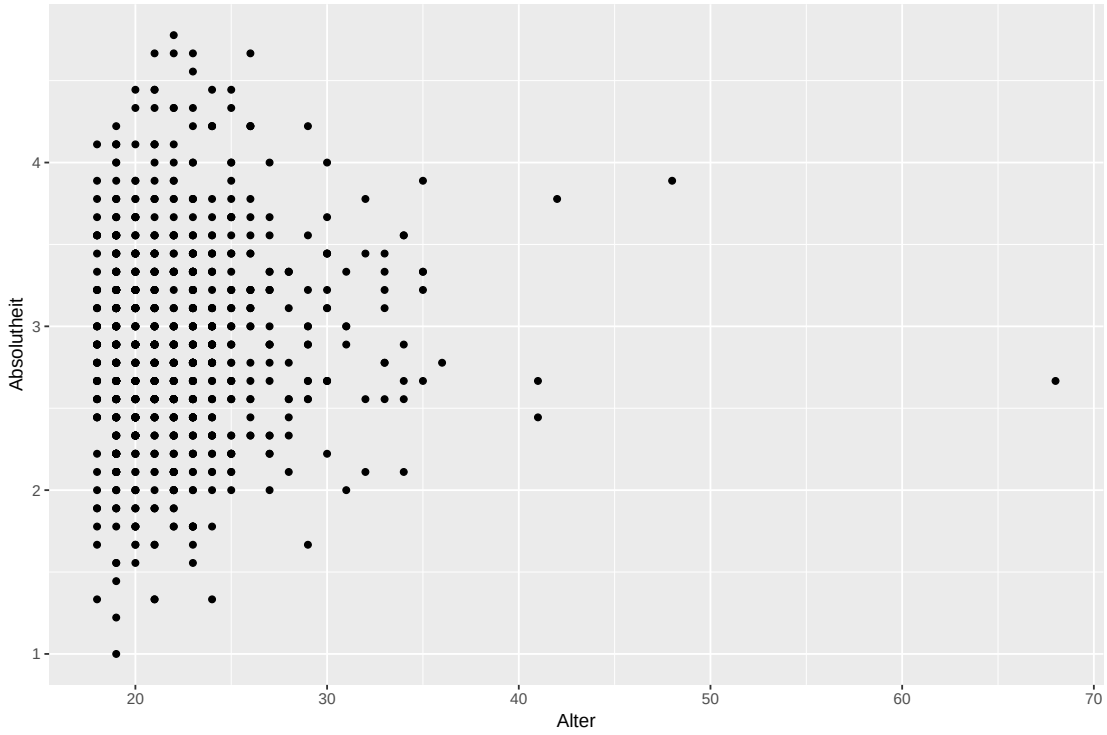


ABBILDUNG 7.1: Streudiagramm für die Überzeugung zur Absolutheit des Wissens und Alter

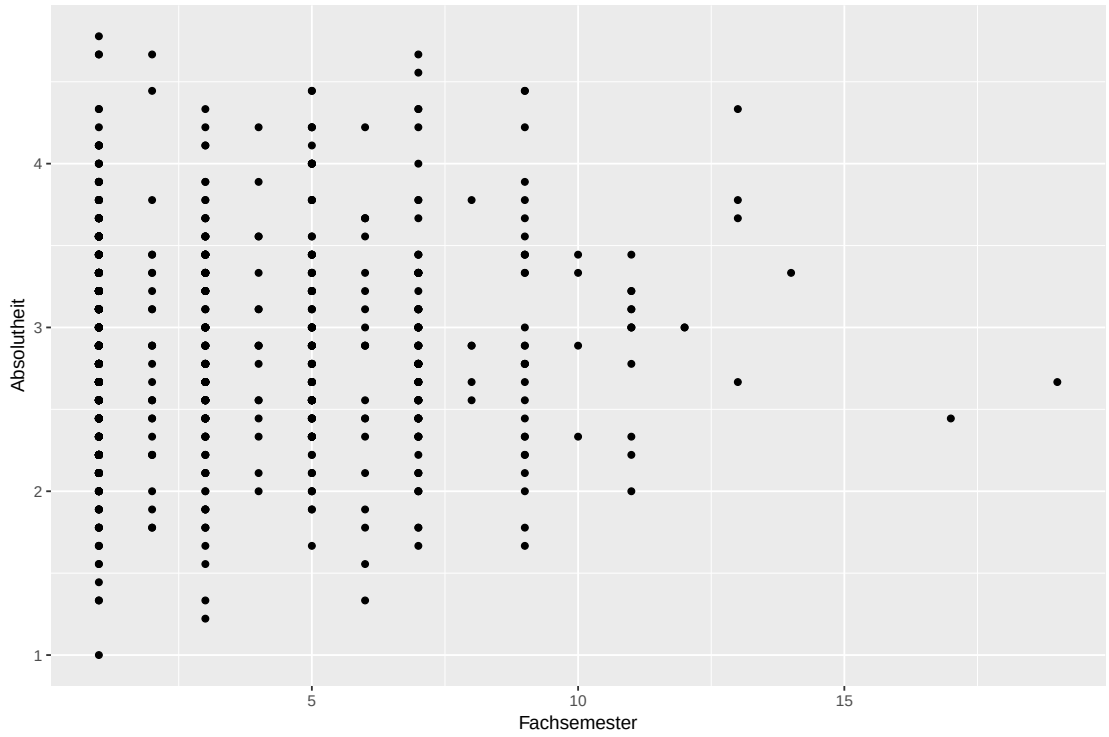


ABBILDUNG 7.2: Streudiagramm für die Überzeugung zur Absolutheit des Wissens und Fachsemester

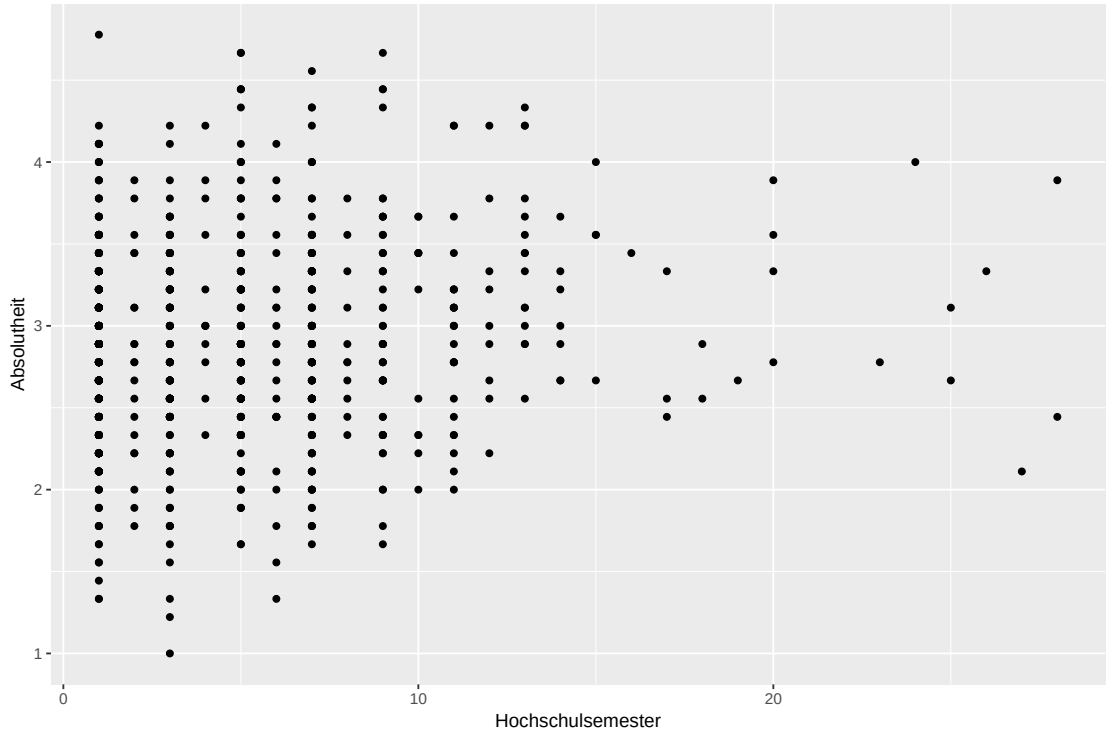


ABBILDUNG 7.3: Streudiagramm für die Überzeugung zur Absolutheit des Wissens und Hochschulsemester

Auf Basis der Analyse der Ausreißer wurden insgesamt acht Datensätze von den weiteren Analysen ausgeschlossen, je zwei aufgrund des abweichenden Alters (> 42) und Anzahl der Fachsemester (> 15) und vier aufgrund des Hochschulsemesters (> 25). Die zu analysierende Stichprobe besteht demnach aus $N$ = 915 Lehramtsstudierenden. Im Kapitel 7.4 (ab S. 221) wurde diese Stichprobe bereits beschrieben.

Den unterschiedlichen Schätzmethoden liegen verschiedene Annahmen zu Skalenniveau und Verteilung zugrunde (vgl. Kap. 4.3.2, S. 160). Die Maximum-Likelihood-Methode, die die am häufigsten eingesetzte CFA-Methode darstellt (Bühner, 2011, S. 408), setzt multivariate Normalverteilung (MVN) und mindestens Intervallskalenniveau voraus (Bühner, 2011, S. 431 f.). Auch die GLS-Methode setzt diese Annahmen voraus. Gerade bei Daten, die mittels Likert-Items erhoben wurden, die ein zugrundeliegendes kontinulierliches und normalverteiltes Konstrukt annehmen, liegt jedoch häufig keine Normalverteilung vor (Schmitt, 2011, S. 307). Entsprechen die empirischen Eigenschaften der Daten nicht den zugrundeliegenden Annahmen der eingesetzten ML-Schätzmethode, kann dies zu verzerrten Parameterschätzungen und Standardfehlern führen (Schmitt, 2011, S. 307). Das führt zu einer Überschätzung des $\chi^2$-Wertes und Verkleinerung des p-Wertes, was unter Umständen zur Verwerfung eines eigentlich geeigneten Modells führen kann (Enders, 2010, S. 140). Die ADF/WLS-Methode (Asymptotically Distribution-Free) benötigt keine Verteilungsannahmen und ist bei ordinalen, dichotomen oder kontinuierlichen manifesten Variablen im Modell geeignet. Es gibt jedoch Hinweise darauf, dass die ADF-Methode nur bei $N > 1000$ zu besseren Parameterschätzungen und genaueren Schätzungen des Modell-Fits als die ML-Methode führt (Olsson et al., 2000). Die ML-Methode ist nach Bühner (2011, S. 409, S. 413) allen anderen Methoden vorzuziehen, da die Parameterschätzungen relativ robust gegenüber Verletzung der multivariaten Normalverteilung (MVN) seien. So postulieren West et al. (1995) Grenzen für die Schiefe und den Exzess der Items innerhalb derer die Durchführung der ML-Methode trotz Abweichung von Normalverteilung noch erlaubt ist (vgl. Bühner, 2011, S. 448). Liegt für alle Items jeweils die Schiefe unter 2 und der Exzess unter 7 darf die ML-Methode trotz Verletzung der multivariaten Normalverteilung angewendet werden, wenn der p-Wert und die Standardfehler mittels beispielsweise Bollen-Stine-Bootstrap (Bollen & Long, 1993) angepasst werden, oder aber eine spezielle robuste ML-Variante gewählt wird, die diese Anpassungen vornimmt. Die MLR-Methode (Maximum-Likelihood robust) bietet gleichzeitig den Vorteil, dass sie sich nicht nur für intervallskalierte (kontinuierliche) Daten eignet, die nicht normalverteilt sind, sondern explizit auch für Rating-Items (Schmitt, 2011, S. 307).

Beim Mardia-Test (Mardia, 1970) auf multivariate Normalverteilung handelt es sich um einen Chi-Test, der abhängig von der Stichprobengröße ist (Brown, 2015). Bei der vorliegenden großen Stichprobe von $N$ = 915, ist davon auszugehen, dass bereits kleine Abweichungen zu einem signifikanten p-Wert und somit zu einer formalen Ablehnung des Modells führen. Mit einem Ergebnis von $p < .001$ zeigt der Mardia Test wie erwartet eine Verletzung der multivariaten Normalverteilung für die StEB-Items. Die StEB-Items liegen jedoch mit Werten für die Schiefe von -0.87 bis 0.59 und Exzess von -0.81 bis 0.57 nahe 0 (siehe Tab. 7.9

S. 233) und deutlich innerhalb der von West et al. (1995) vorgeschlagenen Grenzen (Schiefe < 2, Exzess < 7). Die ML-Methode kann somit trotz der Verletzung der multivariaten Normalverteilung angewandt werden. Es wird aufgrund des Likert-Formats der manifesten Variablen und der formalen Verletztung der Normalverteilungsannahme der robuste ML-Schätzer (MLR) gewählt, der im Lavaan Paket von R (Rosseel, 2012) als Schätzmethode verfügbar ist.

TABELLE 7.9: Schiefe und Exzess der StEB-Items

| Items | *n* | skew | kurtosis |
|---|---|---|---|
| gw_17 | 914 | -0.85 | 0.37 |
| rw_33_u | 915 | 0.21 | -0.15 |
| gw_3_u | 915 | 0.45 | -0.46 |
| ew_5_u | 914 | 0.13 | -0.34 |
| rw_11 | 915 | -0.79 | 0.19 |
| sw_2_u | 913 | 0.63 | -0.15 |
| qw_21_u | 914 | 0.36 | 0.48 |
| ew_13 | 914 | -0.39 | -0.18 |
| sw_11_u | 915 | 0.28 | -0.43 |
| rw_18_u | 915 | 0.32 | -0.43 |
| rw_27 | 915 | -0.72 | 0.04 |
| ew_8 | 913 | 0.02 | -0.43 |
| rw_8_u | 915 | 0.40 | 0.19 |
| sw_17_u | 913 | 0.58 | -0.14 |
| sw_22 | 915 | -0.60 | 0.35 |
| sw_15_u | 912 | 0.60 | -0.47 |
| rw_12 | 915 | -0.37 | -0.22 |
| ew_3 | 914 | 0.02 | -0.29 |
| qw_25_u | 915 | -0.07 | -0.41 |
| rw_36_u | 912 | 0.03 | -0.73 |
| ew_10 | 912 | -0.22 | -0.43 |
| gw_9 | 914 | -0.59 | -0.23 |
| qw_26_u | 915 | 0.06 | -0.77 |
| ew_12 | 914 | -0.15 | -0.41 |
| qw_11 | 914 | -0.64 | -0.04 |
| sw_3_u | 913 | -0.78 | 0.56 |
| qw_13_u | 915 | -0.11 | -0.40 |
| sw_9 | 913 | -0.49 | 0.19 |
| gw_13_u | 914 | 0.02 | -0.80 |
| ew_2 | 914 | -0.33 | -0.42 |
| gw_4_u | 913 | 0.33 | 0.00 |
| gw_16 | 914 | -0.17 | -0.32 |
| rw_29_u | 914 | 0.25 | -0.52 |
| ew_16_u | 914 | 0.28 | -0.51 |

Sehr hoch miteinander korrelierende Items von $r > | .85 |$ können im Rahmen der CFA zu Schätzproblemen führen. Aus diesem Grund sollte vor der Durchführung der CFA sichergestellt werden, dass sogenannte Kollinearität nicht vorliegt. Eine Inspektion der Itemkorrelationsmatrix (siehe Anhang F Tab. F.1 und F.2, ab S. 420) ergab Korrelationen (Pearson)

von minimal $r$ = -.34 und maximal $r$ = .52. Kollinearität liegt dementsprechend nicht vor, und die Voraussetzung ist erfüllt.

Für die CFA gilt, dass größere Stichproben zu genaueren Parameterschätzungen führen. Mit einer Stichprobengröße von $N$ = 915 (nach der Entfernung der Außreißer) ist die empfohlene minimale Stichprobengröße von 200 bis 250 (Bühner, 2011, S. 432) für die ML-Methode deutlich überschritten. Diese Voraussetzung der CFA ist eindeutig erfüllt. Es ist sogar möglich, die Stichprobe nach Zufallsprinzip in zwei Stichproben zu teilen, um die zweite Hälfte zur Kreuzvalidierung zu nutzen, wenn Modifikationen des Modells dies nötig machen sollten.

Die letzte Voraussetzung betrifft die Anzahl von Indikatoren je latenter Variable. Diese sollte nach Bühner (2011) mindestens vier betragen, „damit sich die Identifizierbarkeit des Modells erhöht"(Bühner, 2011, S. 432). Auswirkungen auf den Modellfit ergeben sich durch die Anzahl der Indikatoren pro latenter Variable nicht. Eine größere Anzahl von Items pro Faktor führt häufiger zu angemessenen Lösungen, exakteren und stabileren Parameterschätzungen, weniger konvergierenden Lösungen und reliableren Faktoren (Marsh et al., 1998, S. 213). Das CFA Modell besteht, abgeleitet aus der exploratorischen Faktorenanalyse der ersten Stichprobe, aus vier Subskalen, die über sieben bis zehn Items verfügen. Auch diese Voraussetzung ist somit voll und ganz erfüllt.

### 7.5.2 Ergebnisse der CFA

Die CFA wurde in R Statistics (Version 3.4.0, Team, 2017) mit dem R Paket Lavaan (Rosseel, 2012) umgesetzt. Nach Elimination der Ausreißer verbleibt eine Stichprobe von $N$ = 915 (weiblich 569, männlich 341, k.A. 5) im Alter von 18 bis 42, mit einem durchschnittlichen Alter von 22 (*SD* = 3.34). Zum Umgang mit fehlenden Werten wird die *full information maximum likelihood Prozedur* (FIML) verwendet. Als Schätzmethode wird der robuste ML-Schätzer (MLR) eingesetzt, um der Tatsache Rechnung zu tragen, dass die StEB-Daten nicht multivariat normalverteilt sind und es sich um Likert-Items handelt. Im Rahmen des Berichts über die Modellpassung (vgl. Kap. 4.3.2, Richtlinien zum Modell-Fit in Tab. 4.3, S. 165) werden im folgenden jeweils ausschließlich die robusten, das heißt die bereits angepassten Fitindizes berichtet ($\chi^2$, $\chi^2/df$, RMSEA (CI90), CFit (RMSEA < .05), SRMR, CFI).

#### Spezifikation des ersten vierfaktoriellen Gesamtmodells

Den Anfang der konfirmatorisch-faktorenanalytischen Auswertungen bildet die Spezifikation des vierfaktoriellen Gesamtmodells mit der Zuordnung der Items zu den Faktoren, wie sie aus der exploratorischen Faktorenanalyse und der anschließenden Rekonstruktion des Instruments resultierte. Das umfasst die vier latenten Variablen Absolutheit (*Absol*), Einfachheit (*Simpl*), Multimodalität des Wissenserwerbs (*Multi*) und Entwicklung (*Devel*) und die zugehörigen 34 manifesten Variablen (siehe Abb. 7.4, S. 235). Es handelt sich um ein rekursives Modell, da die Fehlervarianzen unkorreliert sind und alle Pfade zwischen den latenten Variablen und Indikatoren gerichtet sind (Bühner, 2011, S. 446).

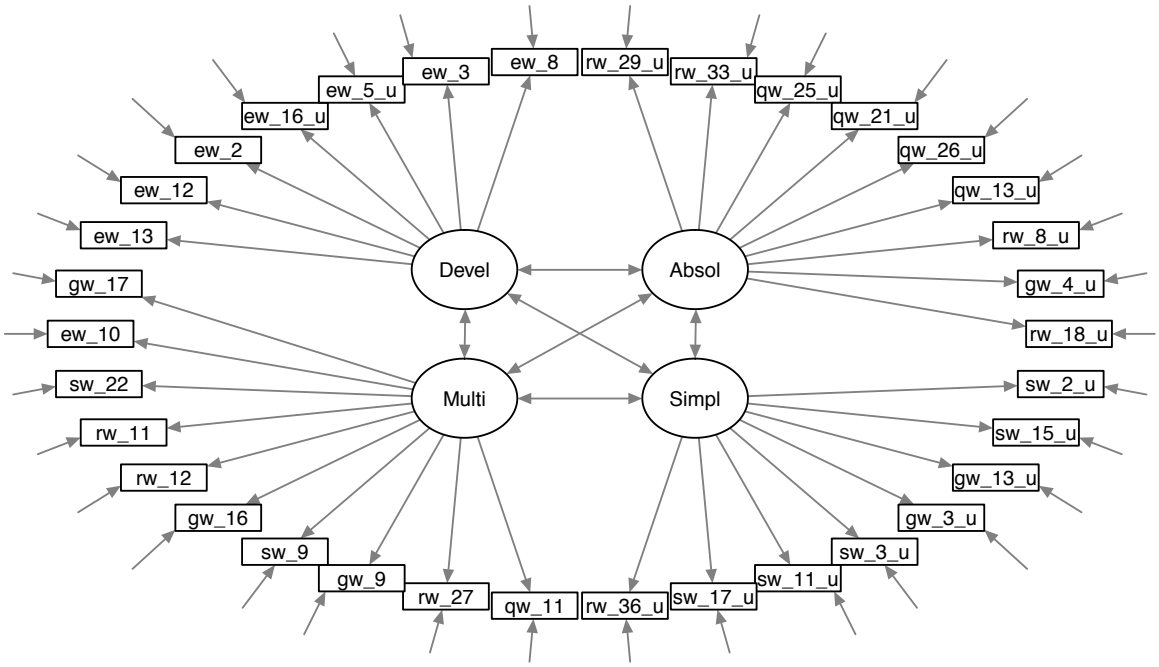


ABBILDUNG 7.4: Rekursives vierfaktorielles Gesamtmodell

Zur Festlegung der Metrik der latenten Variable gibt es zwei Möglichkeiten. Entweder kann die Varianz der latenten Variable auf eins fixiert werden oder es muss die Ladung einer Referenzvariable sowie die Ladung aller Fehlervariablen auf die Items auf eins fixiert werden. Aus der ersten Möglichkeit resuliert eine standardisierte Lösung. Die latente Variable wird standardisiert und alle Faktorladungen können sowohl frei geschätzt als auch auf Signifikanz geprüft werden. Diese Fixierungsvariante ist für die Fragebogenentwicklung sinnvoll, wenn die Höhe der Ladungen der Items noch unbekannt ist (Bühner, 2011, S. 398). Zwar wurden die Ladungen in einer ersten kleineren Stichprobe bereits ermittelt, allerdings handelt es sich um die erste konfirmatorische Faktorenanalyse, der dieses Instrument unterzogen wird. Aus diesem Grund ist es sinnvoll, die Ladungen für alle Items frei zu schätzen und deren Signifikanz zu prüfen. Wird die zweite Fixierungsvariante über eine Referenzvariable gewählt, resuliert eine unstandardisierte Lösung, bei der die latente Variable dieselbe Einheit aufweist wie die Referenzvariable. Die unstandardisierte Lösung ist jedoch nur sinnvoll interpretierbar, wenn alle Items das gleiche Antwortformat aufweisen. Für Likertskalen, wie im vorliegenden Fall, wird diese zweite Variante explizit empfohlen, da alle Items nicht nur das gleiche Format aufweisen, sondern deren Antwortskalenwerte zudem in Einserschritten skaliert sind (Bühner, 2011, S. 397). Eine Interpretation der unstandardisierten Lösung ist im vorliegenden Fall also inhaltlich sinnvoll. Die Bestimmung eines Items als Referenzvariable ist nicht willkürlich. Es sollte dasjenige Item gewählt werden, das am reliabelsten ist. Zur Identifikation der geeigneten Referenzvariablen je Skala sowie zur freien Schätzung aller Ladungen und Bestimmung der Signifikanzen der Ladungen wird zunächst ein vierfaktorielles Gesamtmodell spezifiziert, bei dem die Varianzen der vier latenten Variablen auf eins fixiert werden.

Der globale Modelltest wird mit $\chi^2(524) = 1594.458$ signifikant ($p < .001$). Das Modell wird folglich durch den Modelltest in der vorliegenden Form abgelehnt. Exakter Modell-Fit liegt nicht vor (vgl. Bühner, 2011, S. 422). Dies ist allerdings – trotz der Anwendung der MLR Schätzmethode – aufgrund der starken Abhängigkeit des Modelltest von der Stichprobengröße erwartbar. Allerdings deutet auch das deskriptive Verhältnis des $\chi^2$ mit $\chi^2/df = 3.06$ in Richtung einer Ablehnung des Modells beziehungsweise Verbesserungswürdigkeit, da es einen Wert von über 3 erreicht (Homburg & Giering, 1996, S. 13).

In einem nächsten Schritt werden die Parameterschätzungen analysiert. Wie aus Tabelle 7.10 auf Seite 237 zu entnehmen ist, weisen die hervorgehobenen Items qw_26_u, gw_13_u, qw_11 und ew_12 in ihrer jeweiligen Skala die höchsten geschätzten semipartiellen Regressionsgewichte (Ladungen $\lambda_{i1}$) auf. Durch diese Standardisierung repräsentiert die quadrierte Faktorladung den Anteil der Varianz des Indikators (Items), der durch den Faktor erklärt wird (= Kommunalität) (Brown, 2015, S. 52).

TABELLE 7.10: Geschätzte semipartielle Regressionsgewichte (Ladungen $\lambda_{i1}$) der ersten Modellspezifikation ($N$ = 915)

| Factor regressed on item | $\lambda_{i1}$ | *SE* | *z* | *p* |
|---|---|---|---|---|
| Absol | | | | |
| rw_29_u | .636 | .033 | 19.518 | .000 |
| rw_33_u | .524 | .032 | 16.172 | .000 |
| qw_25_u | .664 | .031 | 21.582 | .000 |
| qw_21_u | .465 | .030 | 15.272 | .000 |
| **qw_26_u** | **.758** | .032 | 23.897 | .000 |
| qw_13_u | .574 | .031 | 18.697 | .000 |
| rw_8_u | .293 | .031 | 9.496 | .000 |
| gw_4_u | .412 | .031 | 13.435 | .000 |
| rw_18_u | .525 | .034 | 15.561 | .000 |
| Simpl | | | | |
| sw_2_u | .612 | .033 | 18.813 | .000 |
| sw_15_u | .496 | .037 | 13.497 | .000 |
| **gw_13_u** | **.683** | .040 | 17.178 | .000 |
| gw_3_u | .404 | .040 | 10.087 | .000 |
| sw_3_u | .237 | .037 | 6.429 | .000 |
| sw_11_u | .274 | .040 | 6.788 | .000 |
| sw_17_u | .618 | .036 | 17.097 | .000 |
| rw_36_u | .574 | .043 | 13.495 | .000 |
| Multi | | | | |
| **qw_11** | **.651** | .031 | 21.154 | .000 |
| rw_27 | .488 | .031 | 15.779 | .000 |
| gw_9 | .429 | .036 | 11.962 | .000 |
| sw_9 | .164 | .036 | 4.507 | .000 |
| gw_16 | .312 | .035 | 8.841 | .000 |
| rw_12 | .457 | .033 | 13.658 | .000 |
| rw_11 | .439 | .034 | 12.787 | .000 |
| sw_22 | .363 | .032 | 11.338 | .000 |
| ew_10 | .353 | .035 | 10.194 | .000 |
| gw_17 | .260 | .029 | 9.103 | .000 |
| Devel | | | | |
| ew_13 | .470 | .030 | 15.624 | .000 |
| **ew_12** | **.510** | .032 | 15.889 | .000 |
| ew_2 | .352 | .041 | 8.516 | .000 |
| ew_16_u | -.372 | .034 | -11.015 | .000 |
| ew_5_u | -.272 | .041 | -6.683 | .000 |
| ew_3 | .479 | .031 | 15.633 | .000 |
| ew_8 | .412 | .036 | 11.589 | .000 |

Das hervorgehobene Item verfügt über die jeweils höchste geschätzte Ladung auf seinem entsprechenden Faktor.

In Abbildung 7.5 wird das spezifizierte Modell mit seinen vollständig standardisierten Ladungen dargestellt. Die Items ew_16_u und ew_5_u (Devel) verfügen berechtigterweise über negative Ladungen, da es sich bei diesen beiden Items um die einzigen umgekehrt formulierten Items ihres Faktors handelt und die umgekehrt formulierten Items für die CFA nicht umkodiert wurden. Die Ladungen lassen sich beispielsweise für Item qw_26_u (Absol) folgendermaßen interpretieren: Eine Veränderung der latenten Variable – das heißt der Merkmalsausprägung *Überzeugung zur Absolutheit des Wissens* – um eine Standardabweichung, geht mit einer durchschnittlichen Veränderung des Items qw_26_u um .69 einer Standardabweichung einher. Da eine vierfaktorielle und keine einfaktorielle Lösung vorliegt, sind die standardisierten Ladungen nicht identisch mit den Korrelationen zwischen dem jeweiligen Item und Faktor (Bühner, 2011, S. 452).

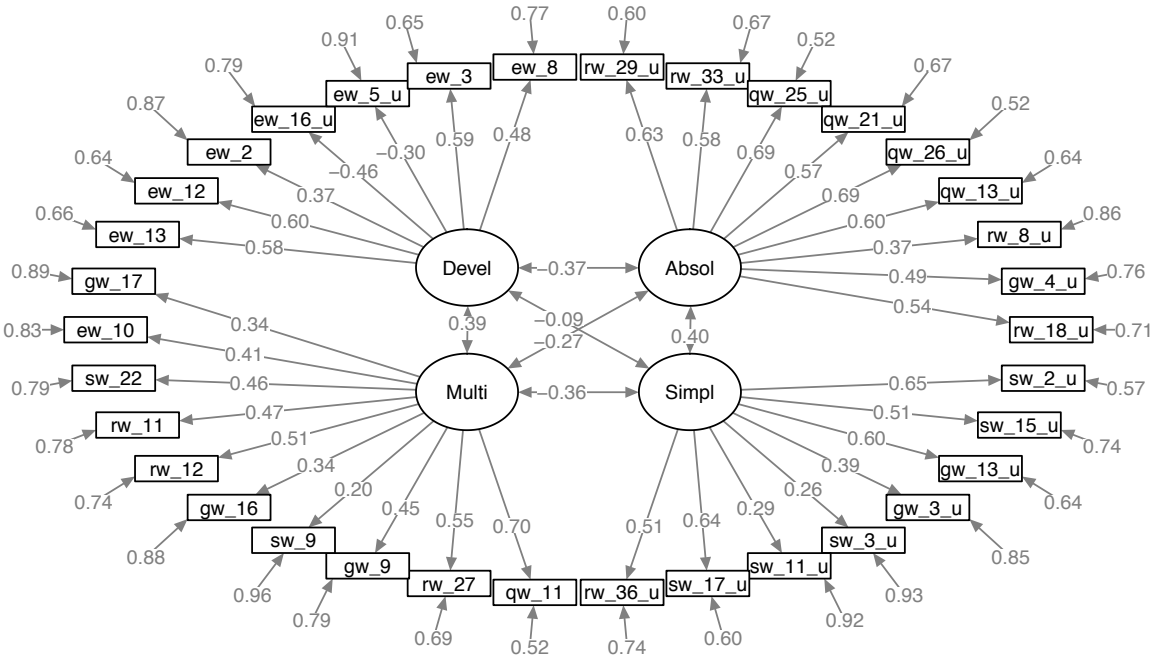


ABBILDUNG 7.5: Standardisiertes vierfaktorielles Gesamtmodell ($N$ = 915)

Die Ladungen aller Items aller vier Faktoren sind signifikant ($p < .001$, siehe Tabelle 7.10, S. 237), was ein Mindestkriterium für Vorliegen des lokalen Modell-Fit darstellt (Bühner, 2011, S. 451). Die meisten Items laden mit mindestens $\lambda > .40$ deutlich auf ihrem eigenen Faktor. Neun Items weisen jedoch mit $\lambda < .40$ geringe standardisierte Ladungen auf.

Der Tabelle 7.11 auf Seite 239 sind die quadrierten Korrelationen der latenten Variable mit den jeweiligen Items zu entnehmen ($R^2$). Dabei handelt es sich um den Varianzanteil des jeweiligen Items, der durch die latente Variable erklärt wird (Bühner, 2011, S. 453). Diese Werte sind als Kommunalitäten, das heißt als Mindestschätzungen der Itemreliabilität zu interpretieren (ebd.). Die vier oben genannten Items erreichen in Bezug zu ihrer jeweiligen Skala die höchsten Kommunalitäten. Inhaltlich bedeutet das, dass diese Items durch die jeweilige latente Variable besser vorhergesagt werden, als die übrigen Items. Folglich eignen sich diese vier Items am besten als jeweilige Referenzvariable: Item qw_26_u für den Faktor

*Absolutheit des Wissens*, gw_13_u für *Einfachheit des Wissens* und qw_11 für *Multimodalität des Wissenserwerbs* und ew_12 für *Entwicklung des Wissens*. Bei sechs Items deuten die sehr niedrigen Kommunalitäten von $R^2 < .13$ und damit einhergehenden sehr hohen Fehlervarianzen $(1 - R^2)$ auf Probleme hin.

TABELLE 7.11: Geschätzte quadrierte Korrelationen der Faktoren mit Items der ersten Modellspezifikation ($N$ = 915)

| Items | $R^2$ | Items | $R^2$ |
|---|---|---|---|
| rw_29_u | .400 | qw_11 | .483 |
| rw_33_u | .331 | rw_27 | .307 |
| qw_25_u | .482 | gw_9 | .206 |
| qw_21_u | .330 | sw_9 | .040 |
| qw_26_u | .480 | gw_16 | .117 |
| qw_13_u | .363 | rw_12 | .264 |
| rw_8_u | .136 | rw_11 | .217 |
| gw_4_u | .242 | sw_22 | .210 |
| rw_18_u | .289 | ew_10 | .168 |
| sw_2_u | .426 | gw_17 | .114 |
| sw_15_u | .264 | ew_13 | .337 |
| gw_13_u | .361 | ew_12 | .360 |
| gw_3_u | .154 | ew_2 | .135 |
| sw_3_u | .066 | ew_16_u | .208 |
| sw_11_u | .082 | ew_5_u | .088 |
| sw_17_u | .405 | ew_3 | .348 |
| rw_36_u | .259 | ew_8 | .231 |

Die lokalen Fit-Indizes RMSEA, CFit und SRMR deuten mit RMSEA = .047 (CI90: .045; .050), CFit = .951 (RMSEA $< .05$) und SRMR = .067 auf akzeptablen bis guten Modell-Fit hin. Der CFI bleibt allerdings mit .80 unter den empfohlen Cut-off-Werten von .90 bzw. .95 (Homburg & Baumgartner, 1995; Hu & Bentler, 1999) zurück. Da das RMSEA Konfizenzintervall den Wert 0 nicht einschließt, ist – übereinstimmend mit dem Ergebnis des globalen Modelltests – nicht von einem exakten Modell-Fit auszugehen (Bühner, 2011, S. 455). Allerdings liegt auch die obere Grenze des Konfidenzintervalls unter dem Cut-off-Wert von .60 (Hu & Bentler, 1999), was wiederum darauf hinweist, dass das Modell zu den empirischen Daten passt (siehe Richtlinien zum Modell-Fit in Tab. 4.3, S. 165).

### CFA des vierfaktoriellen Gesamtmodells und der einfaktoriellen Modelle

Nachdem im Rahmen der ersten Modellspezifikation das jeweils reliabelste Item je Faktor identifiziert sowie alle Faktorladungen geschätzt und deren Signifikanz geprüft wurde, wird nun ein weiteres vierfaktorielles Gesamtmodell spezifiziert. Wie aus Abbildung 7.6 auf Seite 240 zu entnehmen ist, wurde je Faktor die Ladung der Referenzvariablen qw_26_u, gw_13_u, qw_11 und ew_12 auf 1 fixiert. Die Ladungen der übrigen Items werden frei geschätzt. Die Items rw_8_u, sw_3_u und sw_9 verfügen mit $\lambda < .40$ über geringe geschätzte Ladungen (siehe Abb. 7.6, S. 240).

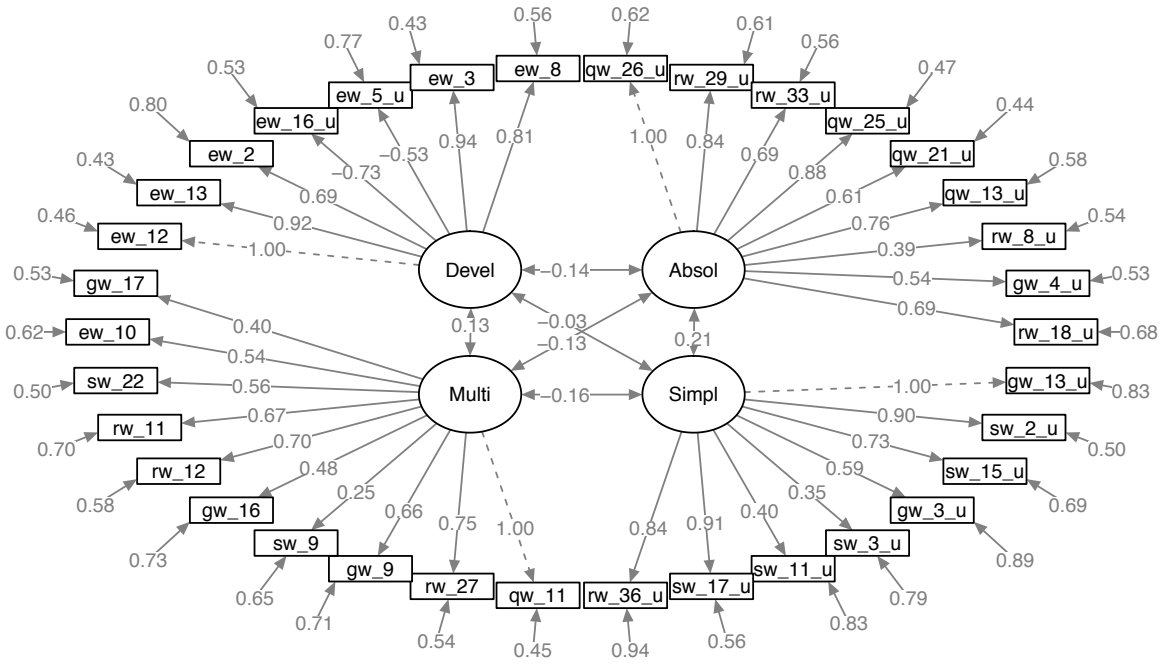


ABBILDUNG 7.6: Geschätztes vierfaktorielles Gesamtmodell ($N$ = 915)

Auch hier sei wieder auf die Items ew_16_u und ew_5_u (Devel) hingewiesen, die berechtigterweise über negative Ladungen verfügen, da es sich bei diesen beiden Items um die einzigen umgekehrt formulierten Items ihres Faktors handelt und die umgekehrt formulierten Items für die CFA nicht umkodiert wurden. Interpretieren lässt sich die unstandardisierte Ladung für beispielsweise Item ew_16_u hier folgendermaßen: Ein Anstieg der latenten Variable um einen Punkt auf der Likertskala (= Metrik der Referenzvariable), geht mit einer durchschnittlichen Veränderung des Items um -0.73 (Abstieg) eines Punktes auf der Likertskala einher. Für die Referenzvariable bedeutet eine Veränderung der latenten Variable um einen Punkt auf der Likertskala ebenfalls eine Veränderung um einen Punkt.

Die geschätzten Varianzen der latenten Variablen sind .574 für *Absolutheit*, .466 für *Einfachheit*, .423 für *Multimodalität des Wissenserwerbs* und .260 für *Entwicklung*. Bedeutend ist, dass auch die Varianzen der vier Faktoren, die in diesem Modell frei geschätzt werden konnten, signifikant sind ($p < .001$). Hierbei handelt es sich um einen Hinweis darauf, dass lokaler Fit des Modells vorliegt (Bühner, 2011, S. 453), da dies aussagt, dass sich die Versuchspersonen in ihren Merkmalsausprägungen – das bedeutet in ihren Überzeugungen zur Absolutheit, Einfachheit, Multimodalität des Wissenserwerbs und Entwicklung – signifikant unterscheiden.

Aus Tabelle 7.12 auf Seite 241 gehen die geschätzen, unstandardisierten Fehlervarianzen der Items, deren Standardfehler und zugehörige z-Werte hervor. Die Fehlervarianz ist die itemspezifische Varianz, die sich nicht durch die latente Variable erklären lässt, während die Standardfehler Aufschluss über den Varianzanteil geben, der sich durch andere Faktoren erklären lässt und der z-Wert ausdrückt, wieviel Varianz unsystematisch ist. Da sich alle Itemfehlervarianzen signifikant von 0 unterscheiden ($p < .001$), lässt sich ein beträchtlicher Varianzanteil der Items über diese drei Möglichkeiten – und eben nicht über die latente

Variable – erklären (vgl. Bühner, 2011, S. 452 f.).

TABELLE 7.12: Geschätzte, unstandardisierte Itemfehlervarianzen ($N$ = 915)

| Items | geschätzte Fehlervarianz | *SE* | *z* | *p* |
|---|---|---|---|---|
| qw_26_u | .623 | .040 | 15.668 | .000 |
| rw_29_u | .607 | .034 | 17.618 | .000 |
| rw_33_u | .556 | .031 | 17.944 | .000 |
| qw_25_u | .474 | .029 | 16.073 | .000 |
| qw_21_u | .440 | .025 | 17.966 | .000 |
| qw_13_u | .579 | .032 | 18.132 | .000 |
| rw_8_u | .545 | .028 | 19.132 | .000 |
| gw_4_u | .532 | .026 | 20.730 | .000 |
| rw_18_u | .678 | .034 | 19.920 | .000 |
| gw_13_u | .826 | .049 | 16.880 | .000 |
| sw_2_u | .505 | .034 | 14.949 | .000 |
| sw_15_u | .687 | .036 | 19.186 | .000 |
| gw_3_u | .893 | .043 | 21.010 | .000 |
| sw_3_u | .794 | .042 | 19.065 | .000 |
| sw_11_u | .834 | .039 | 21.296 | .000 |
| sw_17_u | .562 | .038 | 14.710 | .000 |
| rw_36_u | .943 | .048 | 19.505 | .000 |
| qw_11 | .452 | .035 | 13.109 | .000 |
| rw_27 | .537 | .035 | 15.528 | .000 |
| gw_9 | .712 | .035 | 20.075 | .000 |
| sw_9 | .646 | .033 | 19.727 | .000 |
| gw_16 | .732 | .035 | 20.797 | .000 |
| rw_12 | .580 | .031 | 18.917 | .000 |
| rw_11 | .695 | .042 | 16.721 | .000 |
| sw_22 | .496 | .028 | 17.936 | .000 |
| ew_10 | .615 | .032 | 19.380 | .000 |
| gw_17 | .525 | .028 | 18.695 | .000 |
| ew_12 | .462 | .027 | 17.013 | .000 |
| ew_13 | .434 | .025 | 17.064 | .000 |
| ew_2 | .798 | .041 | 19.474 | .000 |
| ew_16_u | .527 | .028 | 18.511 | .000 |
| ew_5_u | .769 | .036 | 21.170 | .000 |
| ew_3 | .429 | .030 | 14.285 | .000 |
| ew_8 | .565 | .032 | 17.829 | .000 |

Wie im Rahmen der ersten Modellspezifikation bereits erläutert wurde, liegen bei sechs Items sehr niedrige Kommunalitäten vor, was zeigt, das diese Items durch die latente Variable nicht gut vorhergesagt werden. Es scheint sich also in Bezug auf die Erfassung der entsprechenden Dimensionen epistemologischer Überzeugungen um eher ungeeignete Items zu handeln. Aufgrund der Tatsache, dass es sich um dasselbe spezifizierte Modell handelt wie bei der ersten Modellspezifikation, kommen der globale Modelltest sowie die lokalen Fit-Indizes zu identischen Ergebnissen. Der RMSEA liegt mit .047 (CI90: .045;.050) innerhalb der empfohlenen Grenze von RMSEA < .06 (CFit = .951, RMSEA < .05) bei $N > 250$. Auch

der SRMR liegt mit .067 unter dem Cut-off- Wert von < .11 (Hu & Bentler, 1999). Lokaler Modell-Fit ist allerdings nach dem TLI = .79 sowie dem CFI = .803 nicht gegeben und auch globaler Modell-Fit wurde nicht erreicht ($\chi^2(524) = 1594.458$, $p < .001$, $\chi^2/df = 3.06$). Mögliche Gründe für einen eingeschränkten Modell-Fit können nach Brown (2015, S. 147) sein:

1. Die falsche Anzahl von Faktoren wurde spezifiziert.
2. Es liegen Doppelladungen von Items vor, die im Modell nur einem Faktor zugeordnet sind.
3. Items wurden dem falschen Faktor zugeordnet
4. Es liegen korrelierte Fehlervarianzen vor.

Die geschätzte standardisierte Residualkovarianzmatrix der Items, die die inhaltliche Differenz aus beobachteter und modellimplizierter Kovarianz, geteilt durch den jeweiligen Standardfehler, darstellt (Bühner, 2011, S. 459), gibt Hinweise auf für den Modell-Fit problematische Items. Wie der geschätzten, standardisierten Residualkovarianzmatrix im Anhang F in Tabelle F.7 und F.8 (ab S. 427) zu entnehmen ist, liegen für einige Itempaare signifikante Residuen von $-1.96 \leq z \leq 1.96$ ($\alpha = 5\%$) vor. Auch als stark zu klassifizierende Abweichung von $-2.58 \leq z \leq 2.58$, ($\alpha = 1\%$) liegen vor (siehe fett gedruckte Werte in Tab. F.7 und Tab. F.8 (S. 427). Starke Unterschätzungen der beobachteten Kovarianz durch modellimplizierte Kovarianz (pos. z-Werte) häufen sich bei dem auch in anderer Hinsicht problematischen Item ew_5_u (max. $z = 5.460$) mit anderen Items. Die größte Abweichung (Unterschätzung) betrifft die Beziehung zwischen Item ew_10 und ew_8 mit $z = 8.474$. Die Residuen der Items rw_36_u und ew_16_u deuten mit ihren hohen, negativen Werten darauf hin, dass die modellimplizierte Kovarianz die beobachtete Kovarianz stark überschätzt (Bühner, 2011, S. 459). Auch die Items sw_9, ew_2, sw_3_u, rw_8_u, sw_15_u, gw_4_u weisen in ihrer Beziehung zu anderen Items auffällig viele starke Residuen auf.

Um zu prüfen, welche Modellspezifikationen das Modell verbessern würden, lassen sich die Modifikationsindizes heranziehen. Die Modifikationsindizes empfehlen mögliche zu spezifizierende Pfade oder kovariierende Fehlervarianzen, die bisher nicht im Modell umgesetzt sind und deren Spezifikation zu einer Verkleinerung des $\chi^2$-Wertes und demnach Erhöhung des zugehörigen p-Wertes führt. Die angegebene Höhe des Modifikationsindizes enspricht dem Wert, um den sich der $\chi^2$-Wert des Modelltests bei Umsetzung des entsprechenden Pfades verringert. Ab einem Modifikationsindex von 3.84 wird von einer signifikanten Verbesserung des Modell-Fits ausgegangen (Bühner, 2011, S. 454). Für das vorliegende vierfaktorielle Gesamtmodell werden viele Modifikationen vorgeschlagen. Der für dieses Modell vorliegende Modifikationsindex, der mit einem Wert von 90.923 die stärkste Verbesserung des $\chi^2$-Wertes verspricht, empfiehlt, das Item ew_10 dem Faktor Entwicklung zuzuordnen. Allerdings geht daraus nicht hervor, ob das Item ew_10 über eine Doppelladung verfügt und den Faktoren *Entwicklung* und *Multimodalität des Wissenserwerbs* zugeordnet werden sollte oder aber ausschließlich dem Faktor *Entwicklung*. Generell stellt sich

die Frage, ob in der hier spezifizierten Lösung grobe Misspezifikationen vorliegen, die den verringerten Modell-Fit (niedrigen CFI) verursachen und zu einer Ablehnung des Gesamtmodells führen. Solche groben Misspezifikationen wären beispielsweise die Annahme einer falschen Anzahl von Faktoren oder einer falschen Zuordnung der Items zu den Faktoren. Denkbar ist jedoch auch, dass sich nur mehrere kleinere Misspezifikationen aufaddieren und zum verringerten Modell-Fit führen. Kleinere Misspezifikationen bestehen zum Beispiel, wenn von einer Einfachladung eines Items auf einen bestimmten Faktor ausgegangen wird, dieses Item jedoch real über Nebenladungen auf andere Faktoren verfügt (Bühner, 2011).

Um der Frage nach den vorliegenden Misspezifikationen nachzugehen, werden im Folgenden zunächst für jede der latenten Variablen getrennte, einfaktorielle Modelle spezifiziert und geprüft. Dem Anhang D sind in den Abbildungen D.1 bis D.8 auf den Seiten 393 bis 395 die standardisierten und unstandardisierten, geschätzten einfaktoriellen Modelle für die vier Dimensionen zu entnehmen. Aus Tabelle 7.13 geht hervor, dass die vier getrennten einfaktoriellen Modelle im Gegensatz zum vierfaktoriellen Gesamtmodell größtenteils gute Modellpassung erzielen (siehe fett gedruckte Werte).

TABELLE 7.13: Modell-Fit (robust) der vier initialen einfaktoriellen Modelle ($N$ = 915)

| Modell | $\chi^2$ | df | $\chi^2/df$ | $p$ | CFI | RMSEA | CI (90%) | CFit | SRMR |
|---|---|---|---|---|---|---|---|---|---|
| Absol | 78.888 | 27 | **2.92** | .00 | **.97** | **.046** | **.035; .057** | .719* | **.028** |
| Simpl | 70.090 | 20 | 3.50 | .00 | **.94** | **.052** | .040; .065 | .359* | **.035** |
| Multi | 83.144 | 35 | **2.37** | .00 | **.95** | **.038** | **.028; .048** | .972* | **.033** |
| Devel | 56.498 | 14 | 4.04 | .00 | **.94** | **.058** | .043; .073 | .180* | **.035** |

Hervorgehobene Werte befinden sich innerhalb der Cut-off Werte von Hu und Bentler (1999)
* p-close-fit: $.10 \leq p \leq 1.00$ (RMSEA < .05)

Das betrifft hier auch insbesondere den CFI von $\geq .94$. Einzig für die beiden Faktoren Einfachheit und Entwicklung deutet das Verhältnis von $\chi^2/df > 3$ sowie jeweils die rechte Grenze des RMSEA-Konfidenzintervalls von über .06 auf eine Einschränkung der Modellpassung hin. Dies deutet darauf hin, dass im vierfaktoriellen Gesamtmodell notwendige Modellverbesserungen vor allem diese beiden Faktoren zu betreffen scheinen. Gleichzeitig liegt auch nahe, dass bezüglich des Gesamtmodells eher viele kleinere Missspezifikationen für die Einschänkung der Modellpassung verantwortlich sind.

**Rücküberprüfung des spezifizierten Modells mittels EFA**

Ein exploratorischer Ansatz kann genutzt werden, um die Interpretation der durch die CFA implizierten notwendigen Modellmodifikationen zu erleichtern. Indem eine exploratorische Faktorenanalyse durchgeführt wird, die den Items die Zuordnung zu den Faktoren freistellt, lassen sich Unstimmigkeiten in der für das Gesamtmodell spezifizierten Item-Faktor-Zuordnung erkennen.

Zunächst wurde mittels Scree-Plot und Parallelanalyse geprüft, ob es sich bei der richtigen Faktorenanzahl tatsächlich um vier handelt. Dies konnte in beiden Fällen bestätigt werden. Mit einem „verdienstvollen" KMO-Koeffizient (siehe Tab. 4.2, S. 157) von .87 und itemspezifischen MSA-Werten von .74 bis .92 eignen sich die Daten zur Durchführung der Faktorenanalyse. Das Ergebnis der Hauptachsenanalyse mit obliquer Oblimin-Rotation unter Extraktion von vier Faktoren (Pearson Korrelationsmatrix) ist Tabelle 7.14 (S. 245) zu entnehmen. Dargestellt werden die Faktorladungen, Kommunalitäten und die korrigierte Trennschärfe (mit *part-whole-Korrektur*) der Items. Geringe Ladungen unter $\lambda$ = .15 sind aus Gründen der Übersicht nicht dargestellt. Die in zuvorigen CFAs angenommene Item-Faktor-Zuordnung, die sich aus den Analysen der ersten Stichprobe ergeben hat, lässt sich über die Ergebnisse dieser EFA bestätigen. Bis auf eine Ausnahme laden alle Items auf den zuvor angenommenen Faktor. Nur Item ew_10 lädt auf Faktor 4 (*Entwicklung des Wissens*) statt wie auf Basis der ersten Stichprobe angenommen auf Faktor 3 (*Multimodalität des Wissenserwerbs*). Dies entspricht den Ergebnissen der CFA insofern, als dass der Modifikationsindex die Respezifikation des Items ew_10 zu dem Faktor Entwicklung bereits vorgeschlagen hat. Inhaltlich ist diese Veränderung sinnvoll, da es sich bei Item ew_10 um ein Item handelt, dass ganz zu Beginn der Instrumententwicklung tatsächlich für die Entwicklungsskala formuliert wurde.

Die fett gedruckten Werte in Tabelle 7.14 auf Seite 245 stellen in Bezug auf die Höhe der Ladung ($\lambda$ < .40), Kommunalität ($h^2$ < .20) oder Trennschärfe ($r_{ix}$ < .30) geringe Werte dar, die auf problematische Items (sw_3_u, sw_11_u, gw_3_u, gw_9, sw_9, gw_17, ew_2, ew_5_u) hindeuten. Die hier identifizierten Schwächen dieser Items finden ihre Entsprechung in den Ergebnissen der zuvor durchgeführten CFA (vierfaktorielles Gesamtmodell) in Bezug auf $R^2$, standardisierte Itemvarianzen, Ladungen und geschätzte standardisierte Residuen. Für all diese Items hat bereits die CFA sehr geringe ($R^2$ < .13), beziehungsweise im Fall von ew_2 und gw_3_u geringe ($R^2$ < .20), $R^2$ gefunden. Außerdem zeigten diese Items in der CFA bereits hohe Fehlervarianzen und/oder geringe Ladungen ($\lambda$ < .40) (siehe Abb. 7.6, S. 240). Das bedeutet, dass auch durch die Ergebnisse der CFA gezeigt werden konnte, dass es sich um ungeeignete Items handelt, die nur unzureichend von ihrem jeweiligen Faktor erklärt werden können. Auch die zuvor im Rahmen der CFA vorgeschlagenen Modifikationen lassen sich über den Vergleich mit den dargestellten EFA-Ergebnissen einfacher interpretieren. Für die Items, die hier (siehe Tab. 7.14, S. 245) Nebenladungen aufweisen, werden über die Modifikationsindizes der CFA Pfade von den Items zu den entsprechenden Faktoren vorgeschlagen.

Allerdings ist durch den Vergleich der EFA mit den CFA-Ergebnissen nun deutlich geworden, dass es sich hierbei nur um Nebenladungen handelt und nicht um Items, die generell auf einen ganz anderen Faktor als angenommen laden (mit der Ausnahme von Item ew_10). Die in der EFA gefundenen Nebenladungen der Items drücken sich auch über niedrigere standardisierte Ladungen und höhere standardisierte Itemvarianzen in den Ergebnissen der CFA aus (siehe Anhang D in Abb. D.9, S. 396) sowie in hohen geschätzten, standardisierten Residuen dieser Items in Beziehung mit anderen Items (siehe Anhang F in Tab. F.7 und F.8, ab S. 428).

TABELLE 7.14: Mustermatrix der vierfaktoriellen Hauptachsenanalyse mit oblimin-Rotation, Itemkommunalitäten $h^2$ und Itemtrennschärfen $r_{ix}$ ($N$ = 915)

| Items | Faktor 1 | Faktor 2 | Faktor 3 | Faktor 4 | $h^2$ | $r_{ix}$ |
|---|---|---|---|---|---|---|
| rw_29_u | .67 | | | | .43 | .59 |
| qw_26_u | .65 | | | | .46 | .61 |
| qw_25_u | .63 | | | | .46 | .61 |
| qw_21_u | .56 | | | | .34 | .52 |
| rw_18_u | .53 | | | | .31 | .49 |
| rw_33_u | .53 | .15 | .19 | | .39 | .50 |
| gw_4_u | .52 | | | | .26 | .46 |
| rw_8_u | .52 | -.19 | | | .24 | .36 |
| qw_13_u | .48 | .25 | | | .38 | .52 |
| sw_2_u | | .62 | | | .39 | .53 |
| sw_17_u | | .60 | | | .39 | .51 |
| gw_13_u | | .56 | | | .39 | .49 |
| sw_15_u | | .56 | | -.15 | .34 | .45 |
| rw_36_u | .21 | .42 | | | .28 | .39 |
| sw_3_u | -.25 | .42 | | | **.18** | **.26** |
| sw_11_u | | **.34** | | | **.11** | **.26** |
| gw_3_u | .21 | **.31** | | | **.17** | .31 |
| qw_11 | | | .64 | | .46 | .55 |
| rw_12 | | | .55 | | .28 | .42 |
| sw_22 | | | .5 | | .25 | .42 |
| rw_27 | | | .48 | | .32 | .43 |
| rw_11 | | | .42 | | .21 | .37 |
| gw_16 | | | .40 | | **.15** | .32 |
| gw_9 | | | **.37** | | .22 | .36 |
| sw_9 | | -.25 | **.34** | | **.14** | **.21** |
| gw_17 | -.18 | .23 | **.25** | | **.16** | **.27** |
| ew_13 | | | | .60 | .35 | .47 |
| ew_3 | | | | .58 | .37 | .47 |
| ew_12 | | | | .56 | .34 | .47 |
| ew_8 | | | | .53 | .29 | .42 |
| ew_10 | | | .23 | **.47** | .32 | .41 |
| ew_16_u | .19 | | -.19 | .43 | .27 | .36 |
| ew_2 | | .16 | | **.31** | **.17** | .30 |
| ew_5_u | .19 | .23 | | **.24** | **.18** | **.24** |

Die fett gedruckten Werte stellen in Bezug auf die Höhe der Ladung ($\lambda$ < .40), Kommunalität ($h^2$ < .20) oder Trennschärfe ($r_{ix}$ < .30) geringe Werte dar.

**Modelloptimierung und Kreuzvalidierung**

Die Ergebnisse der CFAs zum vierfaktoriellen Modell und zu den einfaktoriellen Modellen sowie der Vergleich dieser Ergebnisse mit denen einer EFA über die gleiche Validierungsstichprobe zeigen deutlich, dass Modellanpassungen notwendig sind. Modellanpassungen müssen allerdings an einer neuen Stichprobe kreuzvalidiert werden (Bühner, 2011, S. 454, S. 460). Indem das gleiche Modell in zwei Stichproben spezifiziert wird, lassen sich die Ergebnisse vergleichen. So kann die Stabilität der Modelllösung geprüft werden und es lässt sich feststellen, ob umgesetzte Modelländerungen in beiden Stichproben zu vergleichbaren Veränderungen in der Modellpassung führen. Ist das nicht der Fall, deutet das darauf hin, dass spezielle Eigenschaften der einen Stichprobe, die in der anderen Stichprobe nicht vorliegen, die Modellpassung beeinflussen (Bühner, 2011, S. 423, S. 429). Liegt eine ausreichend große Stichprobe vor, kann diese randomisiert in zwei Teilstichproben geteilt werden, um so die Kreuzvalidierung der Modellanpassungen vorzunehmen (Bühner, 2011).

Um eine Kreuzvalidierung durchführen zu können, wurden die um die Ausreißer bereinigten Daten der 915 Befragten randomisiert in zwei Substichproben unterteilt von $n_1$ = 457 (Alter: *M* = 22.02, *SD* = 3.33; 164 männlich, 290 weiblich, 3 keine Angabe) und $n_2$ = 458 (Alter: *M* = 21.99, *SD* = 3.36; 177 männlich, 279 weiblich, 3 keine Angabe). Wie auch bei der Gesamtstichprobe kann für je beide Teilstichproben keine multivariate Normalverteilung angenommen werden (Mardia-Test $p < .001$), allerdings sind die itemspezifische Schiefe (< 2) und Wölbung (< 7) jeweils innerhalb der Grenzen nach West, Finch und Curran (1995), sodass die robuste Maximum Likelihood Schätzmethode eingesetzt werden kann (Bühner, 2011, S. 448). Die Prüfung der Korrelationsmatrizen ergab weder für Teilstichprobe 1 ($r$ = -.36 bis $r$ = .54) noch für Teilstichprobe 2 ($r$ = -.31 bis $r$ = .55) zu hohe Korrelationen von $r > .85$.

Zunächst wurde je Teildatensatz ein erstes standardisiertes vierfaktorielles Gesamtmodell spezifiziert mit Fixierung der Varianzen der latenten Variablen auf 1 und den Mittelwert auf 0 (Substichprobe 1: S1, Substichprobe 2: S2). Der Aufbau dieses Modells entspricht in der Item-Faktor-Zuordnung dem bereits geprüften Gesamtmodell (siehe Abb. 7.6, S. 240) und somit der Struktur des StEB Inventars, die aus der EFA in Studie 1 (siehe Kap. 6.5, ab S. 211) resultierte. Die standardisierten vierfaktoriellen Modelle zu Teilstichprobe 1 und 2, bei denen alle Itemladungen frei geschätzt sind, sind den Abbildungen 7.7 und 7.8 auf Seite 247 zu entnehmen. Auffällig sind für die erste Teilstichprobe die sehr niedrigen geschätzten Ladungen von $\lambda < .30$ der Items: rw_8_u, sw_3_u. sw_11_u, sw_9, gw_16, gw_17, ew_5_u. Diese finden sich, mit Ausnahme des Items gw_16, im Modell der zweiten Teilstichprobe wieder. Alle Ladungen beider standardisierter Modelle beider Teilstichproben sind jedoch signifikant mit $p < .001$ (für Item sw_9 aus Teilstichprobe 1 gilt $p = .007$). Für beide Teildatensätze werden die höchsten Itemladungen je Faktor für die Items qw_26_u, gw_13_u, qw_11 und ew_12 geschätzt (siehe Abb. 7.7 und 7.8, S. 247), die somit im unstandardisierten Faktorenmodell geeignete Referenzitems für ihre jeweiligen Skalen darstellen. Im Anhang D in den Abbildungen D.10 und D.11 (siehe S. 397) sind die unstandardisierten geschätzten vierfaktoriellen Modelle für beide Teilstichproben zu sehen.

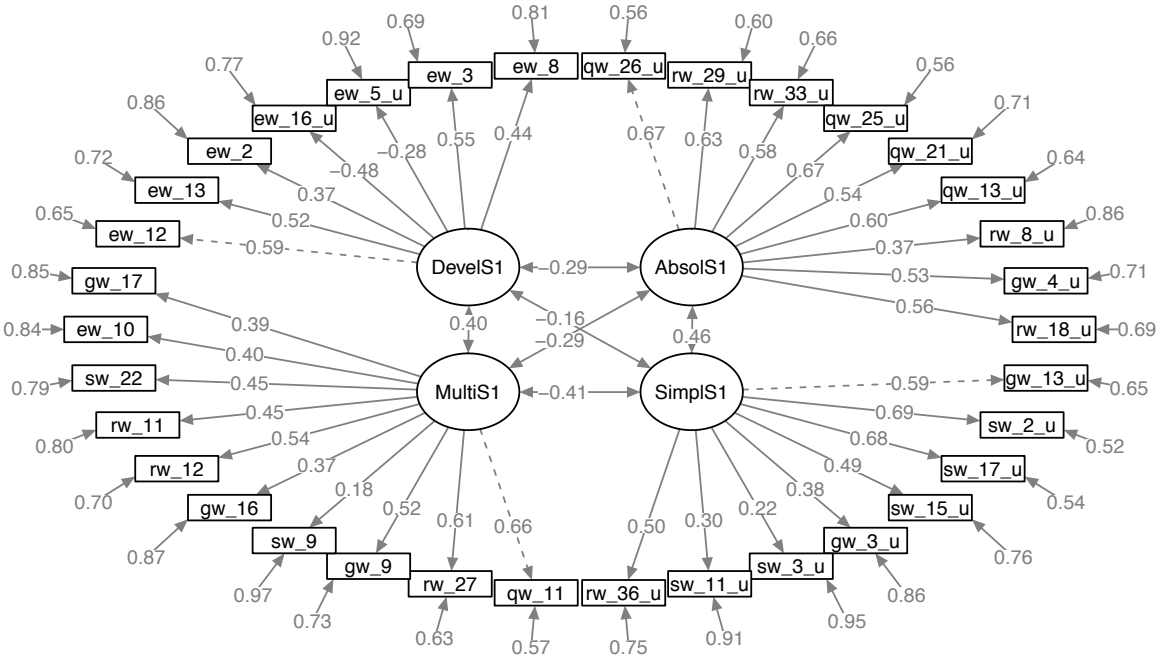


ABBILDUNG 7.7: Standardisiertes vierfaktorielles Modell zu Teilstichprobe S1 ($N$ = 457)

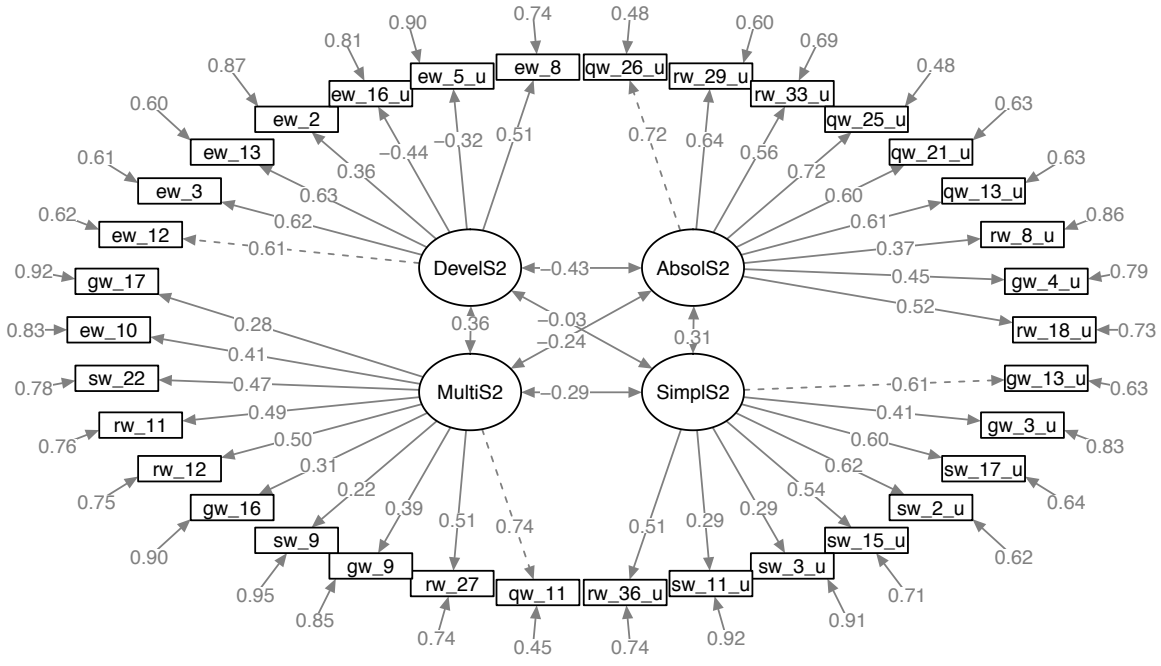


ABBILDUNG 7.8: Standardisiertes vierfaktorielles Modell zu Teilstichprobe S2 ($N$ = 458)

Tabelle 7.15 ist der Modell-Fit der beiden vierfaktoriellen Anfangsmodelle für Teilstichprobe 1 und 2 zu entnehmen. Die fett hervorgehobenen Werte sprechen für eine gute Modellpassung. Zufällige Unterschiede zwischen den Stichprobenteilen sorgen dafür, dass der Modell-Fit der zweiten Teilstichprobe mit dem etwas höheren CFI und niedrigeren RMSEA etwas besser ausfällt. Generell sind die spezifizierten Modelle jedoch vergleichbar.

TABELLE 7.15: Initialer Modell-Fit (robust) der vierfaktoriellen Modelle für Teilstichprobe S1 und S2

| Modelle | $\chi^2$ | df | $\chi^2/df$ | $p$ | CFI | RMSEA | CI (90%) | CFit | SRMR |
|---|---|---|---|---|---|---|---|---|---|
| S1 | 1158.905 | 521 | **2.22** | .000 | .78 | **.052** | **.048; .056** | .221* | **.077** |
| S2 | 1026.484 | 521 | **1.97** | .000 | .82 | **.046** | **.042; .050** | .951* | **.069** |

Hervorgehobene Werte befinden sich innerhalb der Cut-off Werte von Hu und Bentler (1999)
* RMSEA p-close-fit: $.10 \leq p \leq 1.00$ (RMSEA < .05)

Im Folgenden wurde zuerst das unstandardisierte Modell der ersten Teilstichprobe genauer analysiert, um mögliche sinnvolle Modellanpassungen zu identifizieren. Wie schon beim vierfaktoriellen Modell über die gesamte Stichprobe deuten sehr niedrige $R^2(< .13)$, hohe Itemvarianzen und niedrige Ladungen auf problematische Items und somit notwendige Modellanpassungen hin. Das betrifft die oben in Abschnitt 7.5.2 (S. 243) bereits beschriebenen Items. Nach der Empfehlung von Jöreskog (1993) kann die Höhe der Modifikationsindizes als Hinweis für die Reihenfolge der Umsetzung von notwendigen Modellanpassungen genutzt werden. Der höchste der Modifikationsindizes (mi = 48.163) wurde daher zuerst betrachtet. Dieser ordnete das Item ew_10 – übereinstimmend mit den Ergebnissen zur EFA und den zuvor durchgeführten CFAs – dem Faktor Entwicklung zu.

Das Modell der ersten Teilstichprobe wurde entsprechend dieser Änderung respezifiziert und neu berechnet. Anschließend wurde dieser erste Modifikationsschritt auch im Modell der zweiten Teilstichprobe umgesetzt, um die Veränderungen auf Basis der Modellanpassung kreuzzuvalidieren.

Schritt für Schritt wurden weitere Modellanpassungen jeweils zuerst am Modell der ersten Teilstichprobe und im Anschluss am Modell der zweiten Teilstichprobe vorgenommen und die Ergebnisse verglichen. Als Hinweise für vorzunehmende Anpassungen dienten jeweils die Modifikationsindizes des neusten angepassten Modells der ersten Teilstichprobe sowie Hinweise auf problematische Items in Form von niedrigen $R^2$ (< .15), hohen Itemvarianzen (< .85) und niedrigen Ladungen (< .40). Bei allen Modellanpasungen stand die inhaltich sinnvolle Interpretierbarkeit der umzusetzenden Änderungen stets im Vordergrund.

Mittels dieses Vorgehens wurden schrittweise elf Modellanpassungen am vierfaktoriellen Modell der ersten Teilstrichprobe vorgenommen und jeweils am Modell der zweiten Teilstichprobe kreuvalidiert. Die Modell-Fits für diese insgesamt 24 Modelllösungen (berechnete CFAs) sowie die Reihenfolge der umgesetzen Modellanpassungen sind Tabelle 7.16 (S. 249) zu entnehmen. Die erste Anpassung bestand aus der Zuordnung des Items ew_10 zum Entwicklungsfaktor, während in den übrigen zehn Modifikationsschritten nach und nach jeweils ein ungeeignetes Item entfernt wurde.

TABELLE 7.16: Modell-Fit (robust) für die schrittweise modifizierten Modelle der Teilstichprobe S1 und S2

| Schrittweise modifizierte Modelle | $\chi^2$ | df | $\chi^2/df$ | $p$ | CFI | RMSEA | CI (90%) | CFit | SRMR |
|---|---|---|---|---|---|---|---|---|---|
| 4-faktorielles Modell S1 | 1158.905 | 521 | **2.22** | .000 | .78 | **.052** | **.048; .056** | .221* | **.077** |
| 1 Mod S1, ew_10 Wechsel | 1122.721 | 521 | **2.15** | .000 | .79 | **.050** | **.046; .054** | .448* | **.076** |
| 2 Mod S1, sw_9 entfernt | 1051.191 | 489 | **2.15** | .000 | .80 | **.050** | **.046; .054** | .468* | **.074** |
| 3 Mod S1, ew_5_u entfernt | 954.795 | 458 | **2.08** | .000 | .82 | **.049** | **.045; .053** | .689* | **.070** |
| 4 Mod S1, rw_8_u entfernt | 871.615 | 428 | **2.04** | .000 | .83 | **.048** | **.043; .052** | .814* | **.067** |
| 5 Mod S1, ew_2 entfernt | 807.264 | 399 | **2.02** | .000 | .84 | **.047** | **.043; .052** | .834* | **.064** |
| 6 Mod S1, ew_16_u entfernt | 730.504 | 371 | **1.97** | .000 | .85 | **.046** | **.041; .051** | .916* | **.062** |
| 7 Mod S1, gw_17 entfernt | 652.416 | 344 | **1.90** | .000 | .87 | **.044** | **.039; .049** | .972* | **.059** |
| 8 Mod S1, sw_3_u entfernt | 598.225 | 318 | **1.88** | .000 | .878 | **.044** | **.039; .049** | .975* | **.058** |
| 9 Mod S1, sw_11_u | 562.462 | 293 | **1.92** | .000 | .881 | **.045** | **.039; .050** | .943** | **.057** |
| 10 Mod S1, gw_3_u entfernt | 508.608 | 269 | **1.89** | .000 | **.891** | **.044** | **.039; .050** | .958** | **.057** |
| 11 Mod S1, gw 16 entfernt | 476.986 | 246 | **1.94** | .000 | **.892** | **.045** | **.039; .051** | .906* | **.057** |
| 4-faktorielles Modell S2 | 1026.484 | 521 | **1.97** | .000 | .82 | **.046** | **.042; .050** | .951* | **.069** |
| 1 Mod S2, ew_10 Wechsel | 1005.823 | 521 | **1.93** | .000 | .82 | **.045** | **.041; .049** | .979* | **.069** |
| 2 Mod S2, sw_9 entfernt | 945.272 | 489 | **1.93** | .000 | .83 | **.045** | **.041; .049** | .974* | **.068** |
| 3 Mod S2, ew_5_u entfernt | 877.304 | 458 | **1.92** | .000 | .84 | **.045** | **.040; .049** | .980* | **.066** |
| 4 Mod S2, rw_8_u entfernt | 819.054 | 428 | **1.91** | .000 | .85 | **.045** | **.040; .049** | .977* | **.066** |
| 5 Mod S2, ew_2 entfernt | 757.885 | 399 | **1.90** | .000 | .86 | **.044** | **.040; .049** | .980* | **.066** |
| 6 Mod S2, ew_16_u entfernt | 685.8 | 371 | **1.85** | .000 | .87 | **.043** | **.038; .048** | .992* | **.062** |
| 7 Mod S2, gw_17 entfernt | 635.884 | 344 | **1.85** | .000 | .88 | **.043** | **.038; .048** | .990* | **.062** |
| 8 Mod S2, sw_3_u entfernt | 587.604 | 318 | **1.85** | .000 | .883 | **.043** | **.038; .048** | .987* | **.060** |
| 9 Mod S2, sw_11_u | 541.956 | 293 | **1.85** | .000 | **.890** | **.043** | **.038; .048** | .983* | **.060** |
| 10 Mod S2, gw_3_u entfernt | 516.81 | 269 | **1.92** | .000 | .888 | **.045** | **.039; .050** | .935** | **.060** |
| 11 Mod S2, gw 16 entfernt | 495.05 | 246 | **2.01** | .000 | .886 | **.047** | **.041; .053** | .799* | **.061** |

Hervorgehobene Werte befinden sich innerhalb der Cut-off Werte von Hu und Bentler (1999)

* RMSEA p-close-fit: $.10 \leq p \leq 1.00$ (RMSEA < .05)

Die Modelle der beiden Stichproben verhalten sich vergleichbar. Die Modellpassung verbessert sich jeweils für beide Teilstichproben durch jeden Modifikationsschritt bis einschließlich Modifikation 8 (siehe Tab. 7.16, S. 249). Modifikation 9 geht bei Teildatensatz 1 mit einer geringen Verbesserung des CFI und SRMR einher (in der dritten Nachkommastelle), während der RMSEA und der rechte Rand des zugehörigen Konfidenzintervalls um 0.01 steigen. Auch die beiden weiteren Modifikationen sorgen hier für eine Verbesserung des CFI, der RMSEA kann jedoch sein Tief vom Modell mit Modifikation 8 nicht mehr erreichen.

Das Modell der zweiten Teilstichprobe verbessert sich durch den Modifikationsschritt 9 in Bezug auf CFI, beziehungsweise bleibt in Bezug auf RMSEA und SRMR gleich. Die beiden weiteren Modifikationsschritte 10 und 11 bewirken hier jedoch, anders als bei der ersten Teilstichprobe, eine Verschlechterung der Modellpassung in Bezug auf alle Indizes. Die Entfernung von Item gw_3_u und danach Item gw_16 (Modifikationsschritt 10 und 11) wirken sich also in beiden Teilstichproben unterschiedlich auf die Modellpassung aus. Hier scheinen vorliegende unterschiedliche Stichprobeneigenschaften zwischen beiden Teilstichproben wirksam zu werden, die die unterschiedliche Passung beeinflussen. Die letzten beiden Modifikationen lassen sich also nicht mit beiden Stichproben vereinen.

Demnach wird als finales Modell dasjenige gewählt, das dem Modell zu Modifikationsschritt 9 entspricht. Diese werden für beide Teilstichproben in ihrer unstandardisierten Modelllösungen[1] in Abbildung 7.9 und 7.10 (S. 251) gezeigt.

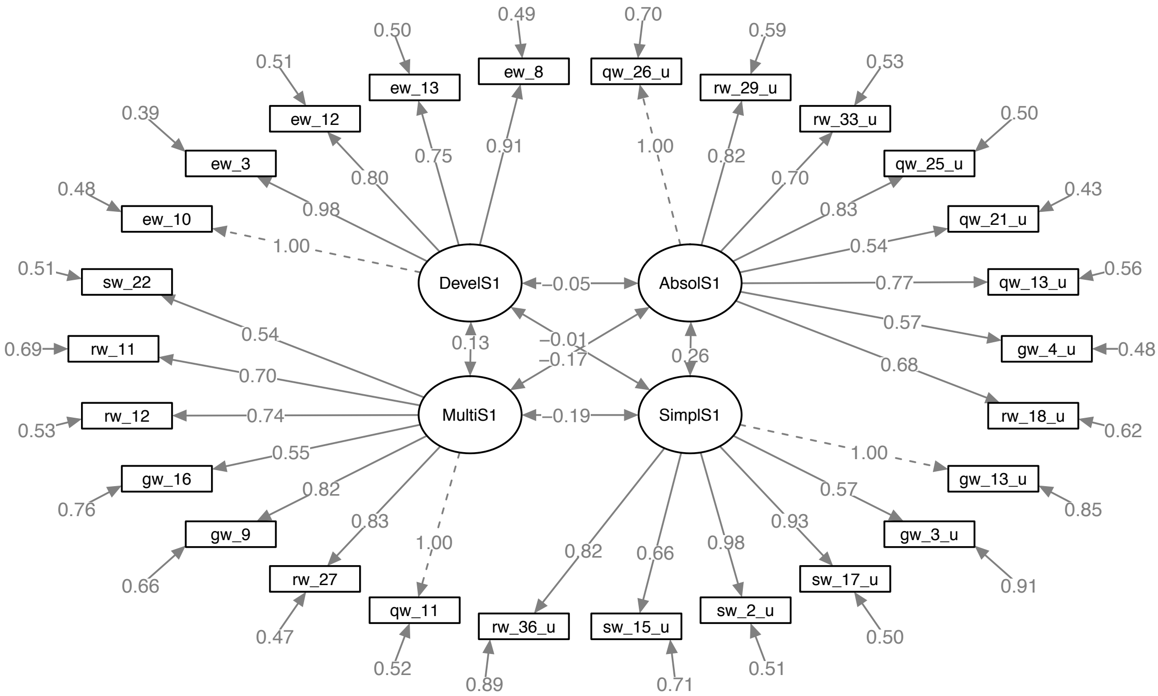


ABBILDUNG 7.9: Finales unstandardisiertes 26-Item Modell zu Teilstichprobe S1 ($N = 457$)

[1] Die standardisierten Modelllösungen sind im Anhang D den Abbildungen D.15 und D.16 auf den Seiten 399, 399 zu entnehmen.

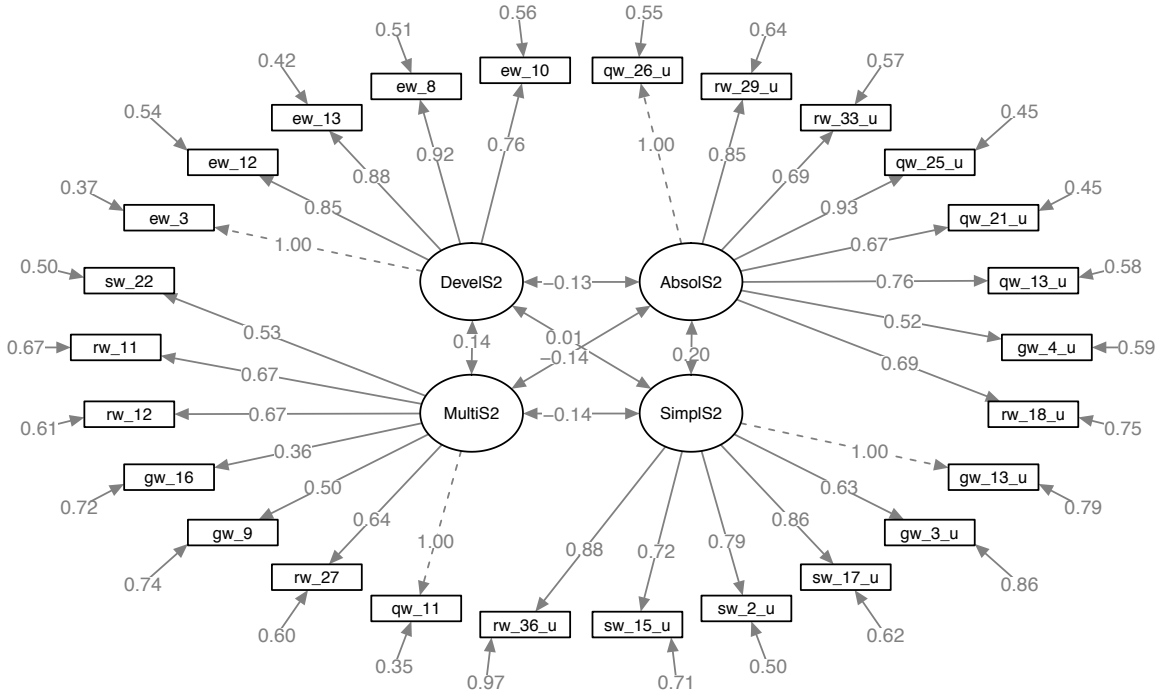


ABBILDUNG 7.10: Finales unstandardisiertes 26-Item Modell zu Teilstichprobe S2 ($N$ = 458)

Die geschätzen Ladungen (semipartielle unstandardisierte Regressionsgewichte) des finalen Modells der Teilstichprobe 1 erreichen Werte von $\lambda$ = .54 bis $\lambda$ = .98. Für die Teilstichprobe 2 werden geschätzte Ladungen von $\lambda$ = .36 bis $\lambda$ = .93 erreicht. Für Item ew_13 (S2) bedeutet die geschätzte Ladung beispielsweise, dass sich der Wert dieses Items um .88 einer Likertstufe ändert, wenn sich die latente Variable, d.h. die Merkmalsausprägung in der Dimension *Entwicklung des Wissens*, um einen Punkt auf der Likertskala ändert. Das bedeutet, dass die Beantwortung der Items kausal auf die vorliegende Merkmalsausprägung auf dem zugrundeliegenden Konstrukt zurückzuführen ist, was für die Validierung der Testwertinterpretationen von entscheidender Bedeutung ist (vgl. Grundannahme 1 in Kap. 3, ab S. 135).

Alle geschätzten Ladungen beider Modelle beider Teilstichproben sind signifikant ($p <$ .001, siehe Anhang E, Tab. E.12, E.17, S. 410, 414). Somit ist ein erstes Mindestkriterium für den lokalen Fit erfüllt (Bühner, 2011, S. 451). Die unstandardisierten Itemvarianzen (Fehlervarianzen) reichen für Teilstichprobe 1 (siehe Abb.7.9, S. 250) von .39 (für Item ew_3) bis .91 (für Item gw_3_u) und für Teilstichprobe 2 (siehe Abb.7.10, S. 251) von .35 (für Item qw_11) bis .97 (für Item rw_36_u). Alle Fehlervarianzen unterscheiden sich signifikant von 0 (siehe Anhang E, Tab. E.16, E.21, S. 413, 417). Das bedeutet, das ein bedeutender Varianzanteil entweder itemspezifisch, unsystematisch oder durch andere Faktoren erklärbar ist. Aufgrund der Modifikationsindizes ist bereits klar, dass ein Teil der Itemvarianz durch andere Faktoren erklärbar ist, da Nebenladungen vorliegen. Die Varianzen der latenten Variablen für beide Stichproben sind alle signifikant (siehe Anhang E, Tab. E.16, E.21, S. 413, 417), das heißt,

dass sich die Personen signifikant in ihren Merkmalsausprägungen bezüglich aller vier Dimensionen unterscheiden. Hierbei handelt es sich um einen Hinweis auf lokalen Modell-Fit (Bühner, 2011, S. 453). Der Anteil der Itemvarianzen, der durch die jeweilige latente Variable erklärt werden kann, bewegt sich für S1 von $R^2$ = .14 (Item gw_16) bis .47 (Item sw_2_u) und für S2 von einem sehr geringen $R^2$ = .08 (Item gw_16) bis .58 (Item qw_11). Eine Übersicht aller $R^2$ den Tabellen E.14 und E.19 (S. 411, 415 im Anhang E) zu entnehmen. Die Residualkovarianzmatrizen zu den beiden Teilstichproben enthalten einige wenige Residuen mit starken Abweichungen ($\mid z \mid \geq 2.58$, zweiseitig, $\alpha$ = 1%) (Byrne, 2010), die Anhaltspunkte für den reduzierten CFI geben. Die erhöhten Residuen treten bei den Teilstichproben in der Beziehung zwischen unterschiedlichen Items auf (siehe Anhang F, Tab. F.9 und F.10, Tab. F.11 und F.10, ab S. 429).

**Finales 26-Item Modell der Gesamtstichprobe**

In einem weiteren Schritt wurde überprüft, wie sich die Modellmodifikationen in der Gesamtstichprobe auswirken. Dazu wurden entsprechend des Modifikationsschritts 9 alle Anpassungen am vierfaktoriellen Modell der Gesamtstichprobe vorgenommen und überprüft. Die Abbildungen 7.11 und 7.12 (S. 253) zeigen die unstandardisierte und standardisierte Modelllösung für das finale 26-Item Modell der Gesamtstichprobe.

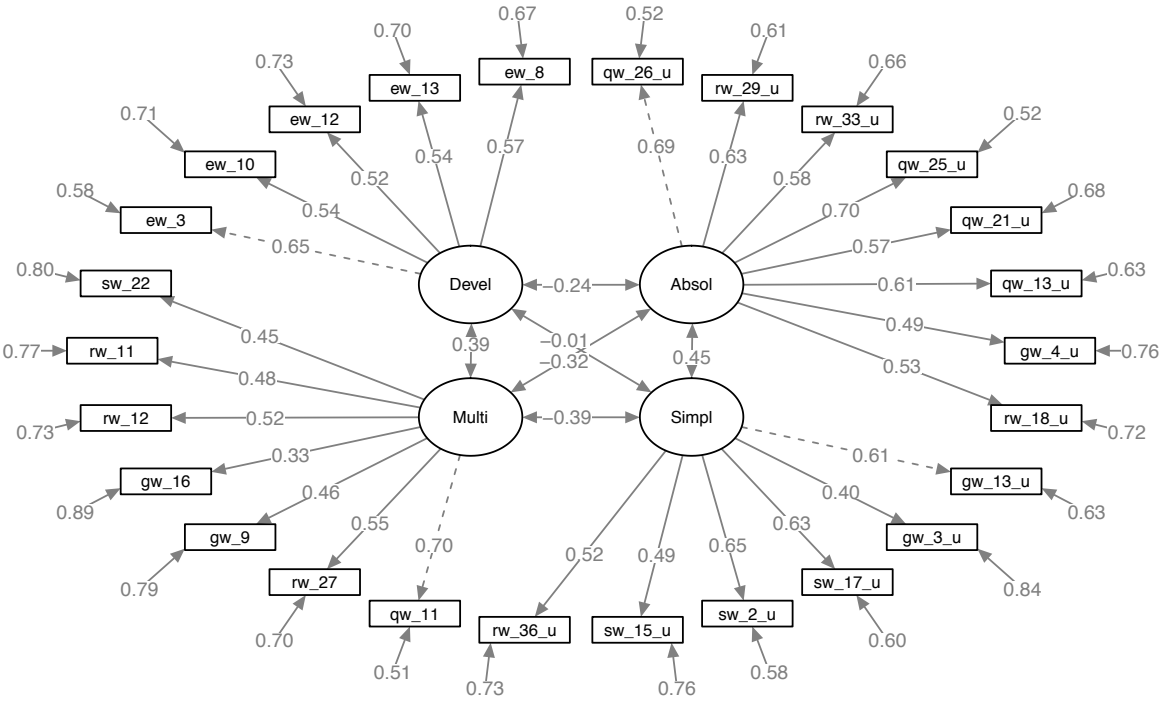


ABBILDUNG 7.11: Finales standardisiertes 26-Item Modell der Gesamtstichprobe ($N$ = 915)

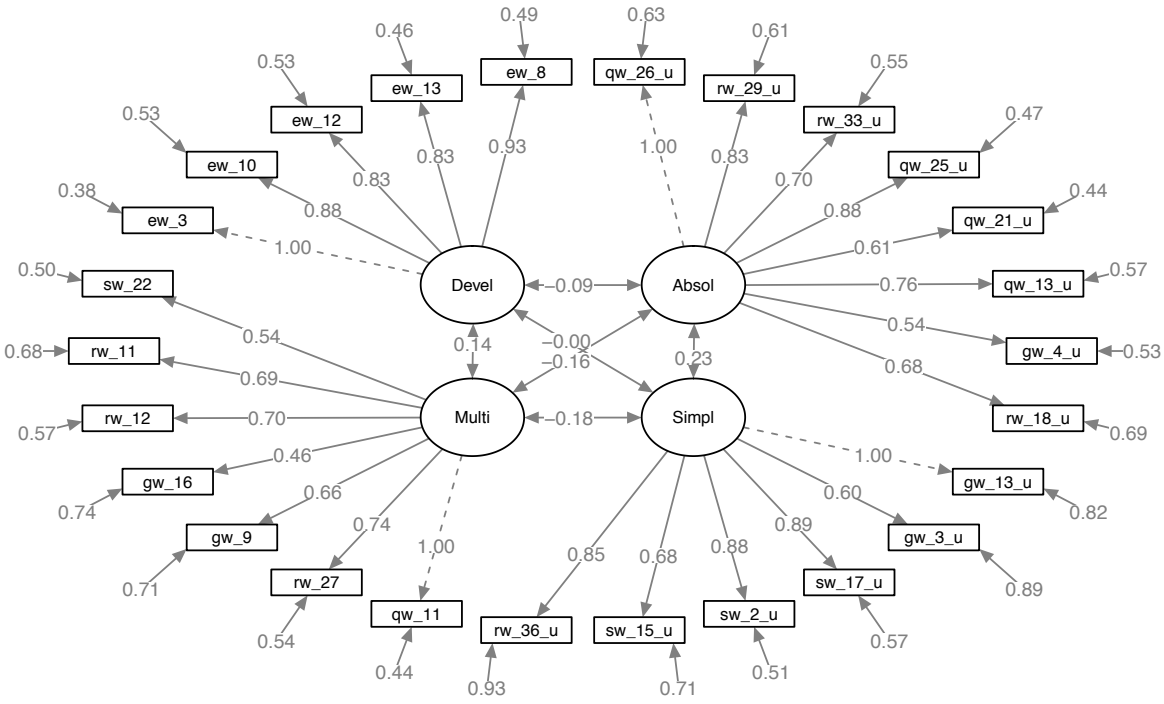


ABBILDUNG 7.12: Finales unstandardisiertes 26-Item Modell der Gesamtstichprobe ($N$ = 915)

Die Parameterschätzungen sind im Anhang E ab Seite 406 einzusehen. Alle geschätzen Itemladungen (von minimal $\lambda$ = .54 für Item sw_22 bis maximal $\lambda$ = .93 für Item ew_8, $p <$ .001), Itemvarianzen und Faktorvarianzen (Absolutheit = .57, Einfachheit = .47, Multimodalität = .43, Entwicklung = .28; $p <$ .001) sind signifikant (siehe Anhang E, Tab. E.7 und E.11, S. E.7, E.11), was ein erstes Mindestkriterium für lokalen Fit darstellt (vgl. Bühner, 2011, S. 451). Der Anteil der Itemvarianzen, der durch die jeweilige latente Variable erklärt werden kann, bewegt sich für das finale 26-Item Modell von $R^2$ = .11 (Item gw_16) bis .49 (Item qw_11). Eine Übersicht der $R^2$ aller Items ist dem Anhang E Tabelle E.9 (S. 407) zu entnehmen.

In Tabelle 7.17 sind die Modell-Fits für die abschließenden vierfaktoriellen 26-Item Gesamtmodelle der Gesamtstichprobe sowie für Teilstichprobe 1 und 2 gegenübergestellt. Die hervorgehobenen Werte entsprechen den Vorgaben nach Hu und Bentler (1999) für gute lokale beziehungsweise globale Modellpassung.

TABELLE 7.17: Modell-Fit (robust) der finalen vierfaktoriellen Modelle der Gesamtstichprobe (GS) und Teilstichprobe S1 und S2

| Modelle | $\chi^2$ | df | $\chi^2/df$ | $p$ | CFI | RMSEA | CI (90%) | CFit | SRMR |
|---|---|---|---|---|---|---|---|---|---|
| GS ($N$ = 915) | 780.837 | 293 | **2.66** | .000 | **.89** | **.043** | **.039; .046** | 1* | **.052** |
| S1 ($N$ = 457) | 562.462 | 293 | **1.92** | .000 | .88 | **.045** | **.039; .050** | .943* | **.057** |
| S2 ($N$ = 458) | 541.956 | 293 | **1.85** | .000 | **.89** | **.043** | **.038; .048** | .983* | **.060** |

Hervorgehobene Werte befinden sich innerhalb der Cut-off Werte von Hu und Bentler (1999)
* RMSEA p-close-fit: $.10 \leq p \leq 1.00$ (RMSEA < .05)

Die globalen Modelltests werden für die drei vierfaktoriellen Gesamtmodelle jeweils mit $p < .001$ signifikant, exakter Modell-Fit kann daher nicht angenommen werden. Allerdings spricht das deskriptive Verhältnis des $\chi^2$ mit $\chi^2/df < 3$ für akzeptablen Fit (Homburg & Giering, 1996, S. 13). Bezüglich der lokalen Modell-Fits liegt RMSEA für alle finalen Modelle nicht nur unter dem Cut-off von .60 (Hu & Bentler, 1999), sondern zeigt mit RMSEA $< .05$ gute Modellpassung (Browne & Cudeck, 1993). Auch liegt die obere Grenze des neunzigprozentigen Konfidenzintervalls des RMSEA unterhalb des vorgegebenen Cut-off-Wertes von .60 (Hu & Bentler, 1999). Darüber hinaus deutet auch der für alle Modelle nicht signifikante CFit (RMSEA CFit < .05) auf akzeptablen Modell-Fit hin. SRMR liegt für alle drei Modelle unterhalb der Grenze von .08 (Hu & Bentler, 1999) und bestätigt demnach den lokalen Fit. Der CFI liegt nahe an der geforderten Mindestgrenze von .90 (Bentler, 1990) nur geringfügig darunter. Auffällig ist, dass der CFI des Modells der ersten Teilstichprobe etwas schlechter ausfällt, als der CFI des Modells der zweiten Teilstichprobe und der des Gesamtmodells für die Gesamtstichprobe (siehe Tab. 7.17, S. 253).

Nach der Prüfung der Modellpassung der vierfaktoriellen Gesamtmodelle wurde überprüft, welche Modellpassung die getrennten einfaktoriellen Modelle erreichen, wenn die oben beschriebenen Modifikationen an den Faktoren vorgenommen werden. Dies wurde sowohl für die Gesamtstichprobe als auch für die beiden Teilstichproben geprüft. Tabelle 7.18 auf Seite 256 zeigt den Modell-Fit für die einfaktoriellen Modellprüfungen der Gesamtstichprobe und der beiden Teilstichproben. Besonders hervorzuheben ist der durchweg sehr gute Modell-Fit des Faktors *Multimodalität des Wissenserwerbs*. Für die beiden Teilstichproben lässt sich aufgrund des nicht signifikanten Modelltests und der Tatsache, dass die untere Grenze des RMSEA-Konfidenzintervalls den Wert 0 annimmt, von exakter Modellpassung sprechen (vgl. Bewertungsrichtlinien für Modell-Fit 4.3, S. 165). Alle Fit-Indizes liegen für die drei Modelle zur *Multimodalität des Wissenserwerbs* unterhalb (RMSEA, RMSEA CI90, SRMR) beziehungsweise überhalb (CFI) der geforderten Cut-off-Werte (Hu & Bentler, 1999). Auffällig sind die Modelle des Faktors *Entwicklung* für die Gesamtstichprobe und die zweite Teilstichprobe. Diese beiden einfaktoriellen Modelle weisen nach den Ergebnissen des globalen Modelltests ($p < .001$, $\chi^2/df > 3$) sowie auch des RMSEA-Wertes (> .06) und dessen Konfidenzintervall (> .06) und CFit (< .10) keinen akzeptablen Modell-Fit auf (vgl. Bewertungsrichtlinien für Modell-Fit 4.3, S. 165). CFI und SRMR liegen wiederum innerhalb der geforderten Werte für guten Fit. Für alle übrigen zwölf Modelle deutet das Verhältnis von $\chi^2$ mit $\chi^2/df < 3$ auf akzeptable bzw. gute $\chi^2/df < 2$ Modellpassung hin. Für die beiden Teilstichproben liegt für je zwei Modelle außerdem – und das trotz der großen Stichproben – ein *nicht* signifikanter globale Modelltest vor ($p > .05$), was ebenfalls auf exakten Fit hindeutet. Für alle zwölf einfaktoriellen Modelle liegt in Bezug auf CFI (> .90) und SRMR (> .08) meist guter bzw. wenigstens akzeptabler Modell-Fit vor (vgl. Bewertungsrichtlinien für Modell-Fit 4.3, S. 165). Der RMSEA liegt für zehn von zwölf Modellen unterhalb des Cut-off-Wertes von .06. Das Konfidenzintervall des RMSEA liegt für den Großteil der einfaktoriellen Modelle allerdings relativ breit um den eigentlichen RMSEA-Wert und mit seinen jeweils oberen Grenzen über .06. Das bedeutet, dass der RMSEA in der Population mit einer gewissen

Wahrscheinlichkeit über dem geforderten Cut-off-Wert von .60 liegt. Wie sich gezeigt hat, liegen auch bei diesen finalen Modellen Einschränkungen in der Passung vor. Dies betrifft vor allem den vierten Faktor, *Entwicklung des Wissens*.

Eine Prüfung der Modifikationsindizes der finalen einfaktoriellen Modelle der Gesamtstichprobe und der Teilstichproben 1 und 2 zeigt, dass die Spezifikation von einigen wenigen korrelierten Fehlervarianzen in den Modellen zu signifikanten Modellverbesserungen führen sollten. Abschließend wurde deshalb überprüft, inwiefern sich exakter Modell-Fit erreichen lässt, wenn korrelierte Fehlervarianzen spezifiziert werden. Ziel war, so wenig wie möglich aber so viele wie nötig korrelierte Itemfehlervarianzen umzusetzen, die über die Modifikationsindizes vorgeschlagen worden sind. Die jeweils erste vorgeschlagene Modifikation des Modells wurde umgesetzt, der Modell-Fit überprüft und bei Bedarf wurde dann die nächste Modifikation umgesetzt und erneut der Modell-Fit geprüft, bis exakter Fit erreicht wurde. Wie sich bezüglich der geschätzen Kovarianzen zwischen den Itemfehlern der einfaktoriellen Modelle (von der Gesamtstichprobe sowie der beiden Teilstichproben) zeigt, sind diese alle sehr gering (nahe 0). Das bedeutet, dass ein geringer oder kein Zusammenhang zwischen den Itemfehlern vorliegt, obwohl die Modifikationsindizes deren Spezifikation vorgeschlagen haben. Die geschätzten Kovarianzen zwischen den spezifizierten, korrelierten Itemfehlern für die vier einfaktoriellen Modelle der Gesamtstichprobe sind im Anhang E in Tabelle E.6 (S.405) einsehbar.

Aus Tabelle 7.19 auf Seite 257 ist zu entnehmen, dass für alle einfaktoriellen Modelle sowohl der Gesamtstichprobe als auch der Teilstichproben exakter Modell-Fit erreicht wird, wenn je nach Modell keine bis maximal drei korrelierte Fehlervarianzen (FV) spezifiziert werden (vgl. Bewertungsrichtlinien für Modell-Fit 4.3, S. 165). Der globale Modelltest wird jeweils nicht signifikant und die untere Grenze des RMSEA Konfidenzintervalls liegt bei .000, während die obere Grenze .60 nicht überschreitet. Die lokalen Fit-Indizes liegen alle deutlich über (CFI, > .90) beziehungsweise unter (RMSEA < .06 und SRMR < .08) den jeweiligen Cut-off-Werten. Dies deutet auf sehr gute Modellpassung hin.

Für das vierfaktorielle Gesamtmodell wird bei Spezifikation von insgesamt fünf korrelierenden Itemfehlervarianzen ein guter Modell-Fit erreicht von: $\chi^2(288) = 682.593$, $p < .001$, $\chi^2/df = 2.37$, RMSEA = .039 (CI90: .035;.042), CFit = 1 (RMSEA < .05), CFI = .91, SRMR = .049. Die geschätzen Kovarianzen zwischen den Itemfehlervarianzen sind jedoch mit .08 bis .13 alle sehr nahe 0, was anzeigt, dass ein geringer oder kein Zusammenhang vorliegt. Die geschätzten Kovarianzen zwischen den spezifizierten, korrelierten Itemfehlern für das vierfaktorielle Modell der Gesamtstichprobe sind in Tabelle E.6 (S. 405) im Anhang E einsehbar. Die übrigen Parameterschätzungen dieses Modells sind ebenfalls dem Anhang E zu entnehmen (siehe Tab. E.1, E.2, E.3, E.4, E.5 ab S. 401).

TABELLE 7.18: Modell-Fit (robust) für die finalen einfaktoriellen Modelle der Gesamtstichprobe und Teilstichprobe S1 und S2

| Finale einfaktorielle Modelle | $\chi^2$ | *df* | $\chi^2/df$ | *p* | CFI | RMSEA | CI (90%) | CFit | SRMR |
|---|---|---|---|---|---|---|---|---|---|
| Gesamtstichprobe ($N$ = 915) | | | | | | | | | |
| 8-Item Absolutheit | 70.746 | 20 | 3.54 | .000 | **.97** | **.053** | .041; .065 | .341* | **.029** |
| 6-Item Einfachheit | 26.282 | 9 | **2.92** | .002 | **.98** | **.046** | .027; .065 | .607* | **.024** |
| 7-Item Multimodalität | 29.799 | 14 | **2.13** | .008 | **.98** | **.035** | **.019; .051** | .935* | **.024** |
| 5-Item Entwicklung | 35.625 | 5 | 7.13 | .000 | **.95** | .082 | .059; .106 | .012 | **.034** |
| Teilstichprobe 1 ($N$ = 457) | | | | | | | | | |
| 8-Item Absolutheit | 42.245 | 20 | **2.11** | .003 | **.97** | **.049** | .029; .069 | .493* | **.031** |
| 6-Item Einfachheit | 17.389 | 9 | **1.93** | .043 | **.98** | **.045** | .011; .075 | .560* | **.028** |
| 7-Item Multimodalität | 18.787 | 14 | **1.34** | **.173** | **.99** | **.027** | **.000; .054** | .915* | **.027** |
| 5-Item Entwicklung | 10.576 | 5 | **2.12** | **.060** | **.98** | **.049** | .002; .088 | .452* | **.027** |
| Teilstichprobe 2 ($N$ = 458) | | | | | | | | | |
| 8-Item Absolutheit | 44.092 | 20 | **2.20** | .001 | **.97** | **.051** | .032; .070 | .427* | **.032** |
| 6-Item Einfachheit | 16.471 | 9 | **1.83** | **.058** | **.98** | **.043** | .003; .073 | .616* | **.027** |
| 7-Item Multimodalität | 21.528 | 14 | **1.54** | **.089** | **.98** | **.034** | **.000; .060** | .826* | **.030** |
| 5-Item Entwicklung | 35.652 | 5 | 7.13 | .000 | **.92** | .116 | .082; .152 | .001 | **.042** |

Hervorgehobene Werte befinden sich innerhalb der Cut-off Werte von Hu und Bentler (1999)
* RMSEA p-close-fit: $.10 \leq p \leq 1.00$ (RMSEA < .05)

TABELLE 7.19: Modell-Fit (robust) für die finalen einfaktoriellen Modelle mit korrelierten Fehlervarianzen der Gesamtstichprobe und Teilstichprobe 1 und 2

| Finale einfaktorielle Modelle mit korr. FV | $\chi^2$ | *df* | $\chi^2/df$ | *p* | CFI | RMSEA | CI (90%) | CFit | SRMR |
|---|---|---|---|---|---|---|---|---|---|
| Gesamtstichprobe ($N$ = 915) | | | | | | | | | |
| 8-Item Absolutheit, 3 korr. FV | 26.316 | 17 | 1.55 | .069 | .994 | .024 | .000; .041 | .997* | .018 |
| 6-Item Einfachheit, 2 korr. FV | 4.417 | 6 | 0.74 | .620 | 1 | .000 | .000; .034 | .997* | .011 |
| 7-Item Multimodalität, 1 korr. FV | 17.813 | 13 | 1.37 | .165 | .993 | .020 | .000; .039 | .997* | .020 |
| 5-Item Entwicklung, 2 korr. FV | 1.711 | 3 | 0.57 | .635 | 1 | .000 | .000; .041 | .980* | .007 |
| Teilstichprobe 1 ($N$ = 457) | | | | | | | | | |
| 8-Item Absolutheit, 2 korr. FV | 25.917 | 18 | 1.44 | .102 | .990 | .031 | .000; .054 | .902* | .025 |
| 6-Item Einfachheit, 1, korr. FV | 5.995 | 8 | 0.75 | .648 | 1 | .000 | .000; .043 | .976* | .018 |
| 7-Item Multimodalität, 0 korr. FV | 18.787 | 14 | 1.34 | .173 | .99 | .027 | .000; .054 | .915* | .027 |
| 5-Item Entwicklung, 2 korr. FV | 1.065 | 3 | | .786 | 1 | .000 | .000; .043 | .967* | .008 |
| Teilstichprobe 2 ($N$ = 458) | | | | | | | | | |
| 8-Item Modell Absolutheit, 2 korr. FV | 22.526 | 18 | 1.25 | .209 | .994 | .023 | .000; .048 | .963* | .024 |
| 6-Item Einfachheit, 1, korr. FV | 9.740 | 8 | 1.22 | .284 | .995 | .022 | .000; .059 | .875* | .020 |
| 7-Item Multimodalität, 0 korr. FV | 21.528 | 14 | 1.54 | .089 | .98 | .034 | .000; .060 | .826* | .030 |
| 5-Item Entwicklung, 3 korr. FV | 0.342 | 2 | 0.17 | .843 | 1 | .000 | .000; .045 | .961* | .005 |

* RMSEA p-close-fit: $.10 \leq p \leq 1.00$ (RMSEA $< .05$)

### 7.5.3 Zusammenfassung und Diskussion der Ergebnisse

Im Rahmen der zweiten Studie wurde die aus der ersten Studie resultierende 34-Item StEB-Version gemeinsam mit dem LIST-R zur Erfassung von Lernstrategien im Studium (Wild, 2011) an einer zweiten, größeren Stichprobe von zunächst 1051 Lehramtsstudierenden sechs deutscher Hochschulen eingesetzt. Nach der Bereinigung der Daten und dem Ausschluss von acht Ausreißern verblieben die Datensätze von $N = 915$ Lehramtsstudierenden. Die folgende Analyse der fehlenden StEB-Werte lehnte die MCAR Annahme ab. Aufgrund der starken Hinweise auf MAR wurde als geeignetes Verfahren zum Umgang mit den fehlenden StEB-Werten FIML gewählt. MLR wurde nach der Prüfung der CFA-Voraussetzungen bezüglich Skalenniveau, Verteilung und Stichprobengröße der ML-Methode als die angemessene Schätzmethode vorgezogen, da sie keine multivariat normalverteilten Daten benötigt und die Tatsache berücksichtigt, dass Ratingitems vorliegen. Die Spezifikation des ersten CFA-Gesamtmodells wurde auf Basis der aus der ersten Studie resultierenden Faktorstruktur vorgenommen. Sowohl die Parameterschätzungen des standardisierten und unstandardisierten Gesamtmodells als auch die Modifikationindizes und die Residualkovarianzmatrix der Items wurden genau untersucht. Um den identifizierten Missspezifikationen nachzugehen, wurden im Anschluss die vier einfaktoriellen CFA-Modelle (ein Modell je Faktor) spezifiziert und geprüft sowie eine exploratorische Prüfung des Gesamtmodells mittels EFA vorgenommen. Die EFA erleichterte die Interpretation der durch die CFA implizierten notwendigen Modellmodifikationen, indem den Items die Zuordnung zu den Faktoren freigestellt wurde und sich so Unstimmigkeiten in der für das Gesamtmodell spezifizierten Item-Faktor-Zuordnung erkennen ließen. Wie sich in der EFA zeigte, konnte die aus der ersten Stichprobe resultierende Faktorstruktur des StEB Inventars weitgehend repliziert werden. Die Ergebnisse der CFAs zum vierfaktoriellen Gesamtmodell und zu den einfaktoriellen Modellen sowie der Vergleich dieser Ergebnisse mit denen der EFA implizierten verschiedene Modellanpassungen. Die Modellanpassungen bestanden jedoch nur – mit der Ausnahme eines Items, das den Faktor wechselte – im Entfernen ungeeigneter Items. Um die Modellanpassungen kreuzvalidieren zu können, d.h. zu prüfen, ob sie zu einer stabilen Modelllösung führen, wurde die gesamte Validierungsstichprobe randomisiert in zwei Teilstichproben geteilt. So wurden Schritt für Schritt elf Modellrevisionen erst an Teilstichprobe 1 und danach an Teilstichprobe 2 vorgenommen. Für jeden Schritt wurden die Veränderungen in den Parameterschätzungen, den Modifikationsindizes und der Modellpassung beider Modelle geprüft und verglichen. So konnte untersucht werden, ob umgesetzte Modelländerungen in beiden Stichproben zu vergleichbaren Veränderungen in der Modellpassung führten. Neun dieser Modifikationen wurden für das finale Modell beibehalten, da diese neun vorgenommenen Modelländerungen sich in beiden Stichprobenteilen gleich auswirkten. Aus Studie 2 resultiert auf diese Weise eine vierfaktorielle 26-Item Modelllösung. Sowohl für die einfaktoriellen Modelle der Gesamtstichprobe als auch die einfaktoriellen Modelle der beiden Teilstichproben konnte *exakter* Modell-Fit erreicht werden, wenn je nach Modell keine bis maximal drei korrelierte Fehlervarianzen (Messfehler) spezifiziert wurden. Für das vierfaktorielle Modell der Gesamtstichprobe wurde bei Spezifikation von insgesamt fünf korrelierten Itemfehlervarianzen guter Modell-Fit erreicht.

Korrelierte Messfehler (Itemfehlervarianzen) treten häufig bei ähnlich formulierten Items auf, also bei Items, die gleiche Wörter enthalten (vgl. Schermelleh-Engel & Wagner, 2008). Sie kommen auch zwischen Items mit gleicher Formulierungsrichtung vor, wenn eine Skala sowohl positiv als auch negativ formulierte Items enthält (ebd.). Korrelierende Messfehler entsprechen Methodenfaktoren im Sinne von systematischen Varianzquellen, die neben der eigentlichen latenten Variable die Itembeantwortung beeinflussen (ebd.). Die Subskalen des StEB enthalten allerdings nur gleichgerichtet formulierte Items, d.h. auf die Formulierungsrichtung können die korrelierten Itemfehler nicht zurückzuführen zu sein. Es wäre möglich, dass die korrelierenden Messfehler im StEB Inventar auf Effekte der Itemformulierung zurückzuführen sind (vgl. DiStefano & Motl, 2009; Schriesheim, Eisenbach & Hill, 1991). Allerdings sind die vorliegenden Zusammenhänge zwischen den Messfehlern der einfaktoriellen Modelle und des vierfaktoriellen Modells der Gesamtstichprobe (sowie der einfaktoriellen Modelle der Teilstichproben) äußerst gering – hauptsächlich < .10 (vgl. Anhang E, Tab. E.6, S. 405) – weshalb eher davon auszugehen ist, dass sie nur aufgrund der großen Stichprobe signifikant wurden. Demnach wird die vierfaktorielle Modelllösung ohne korrelierte Messfehler als abschließende Modelllösung gewählt.

Die Einschränkungen im Modell-Fit des finalen vierfaktoriellen Modells im CFI (siehe Tab. 7.17, S. 253) sind erklärbar über vorhandene Nebenladungen von Items. Einige der Items laden nicht nur auf ihrem intendierten Faktor, sondern zusätzlich auch noch auf einem anderen Faktor. Diese Nebenladungen gehen deutlich aus den in den Modifikationsindizes der CFAs vorgeschlagenen zu spezifizierenden Pfaden hervor, aus den Ergebnissen der EFA (Studie 2) und auch die signifikanten Itemfehlervarianzen zeigen dies. Da Doppelladungen im Rahmen der CFA jedoch nicht zugelassen werden, liegt eine Diskrepanz zwischen den vorliegenden empirischen Daten und dem spezifizierten Modell vor, die den Modell-Fit einschränkt. Der Fragebogen zielt darauf ab, die epistemologischen Überzeugungen von Lehramtsstudierenden in Bezug auf vier Dimensionen zu erfassen. Die Nebenladungen bei den Items sind inhaltlich über die Tatsache erklärbar, dass es sich bei den vier zu erfassenden Dimensionen um Teilaspekte des übergeordneten Konstrukts *epistemologische Überzeugungen von Lehramtsstudierenden zu erziehungswissenschaftlichem Wissen* handelt. Über die vorhandenen signifikanten Zusammenhänge (Kovarianzen) zwischen den vier Faktoren des finalen vierfaktoriellen Modells ist erkennbar, dass die vier Dimensionen des StEB Teilaspekte eines übergreifenden Konstrukts sind und natürlich nicht voneinander unabhängig sind. Auf diese Zusammenhänge wird im Rahmen der Validierung der Testwertinterpretationen in Kapitel 8.2 eingegangen und sie werden dort in Tabelle 8.1 (S. 278) dargestellt.

Die Items gw_3_u und gw_16 verfügen mit $R^2$ = .16 bzw. $R^2$ = .11 (siehe Anhang E Tab. E.9, S. 407) über einen geringen Varianzanteil, der durch die latente Variable erklärt wird (vgl. Bühner, 2011, S. 453). Inhaltlich bedeutet das, dass diese Items durch die jeweilige latente Variable nicht gut vorhergesagt werden. Item gw_3_u verfügt mit .89 zudem über eine hohe geschätze, unstandardisierte Fehlervarianz (standardisierte Fehlervarianz: .84, siehe Anhang E, Tab. E.11, S. 409). Die Fehlervarianz ist die itemspezifische Varianz, die sich nicht durch die latente Variable erklären lässt. Der zugehörige Standardfehler ($SE$

= .042) gibt Aufschluss über den Varianzanteil, der sich durch andere Faktoren erklären lässt, während der z-Wert ($z$ = 21.059) die unsystematische Varianz ausdrückt. Da sich die Itemfehlervarianz dieses Items (wie auch aller anderen) signifikant von 0 unterscheidet ($p$ < .001), lässt sich ein beträchtlicher Varianzanteil der Items über diese drei Möglichkeiten – und eben nicht über die latente Variable – erklären (vgl. Bühner, 2011, S. 452 f.). Item gw_16 verfügt neben dem niedrigen $R^2$ mit $\lambda$ = .33 auch über eine geringe standardisierte Ladung (siehe Abb. 7.11, S. 252) und mit $\lambda$ = .46 auch über eine relativ geringe unstandardisierte Ladung (siehe Abb. 7.12, S. 253). Inhaltlich bedeutet das, dass eine Veränderung in der Überzeugung zur Multimodalität des Wissenserwerbs um eine Standardabweichung nur mit einer durchschnittlichen Veränderung in der Itembeantwortung von .33 einer Standardabweichung einhergeht. Ein Anstieg in der Überzeugung zur Multimodalität des Wissenserwerbs um einen Punkt auf der Likertskala (= Metrik der Referenzvariable) geht mit einer durchschnittlichen Veränderung der Itembeantwortung um .46 (Anstieg) eines Punktes auf der Likertskala einher. Eine Veränderung in der Merkmalsausprägung (Überzeugung zur Multimodalität des Wissenserwerbs) hat also einen relativ geringen Einfluss auf die Veränderung in der Itembeantwortung. Aus diesem Grund wurde in der Modellanpassung das Enfernen dieser beiden Items in Betracht gezogen (siehe Modifikationsschritte 10 und 11 in Tab. 7.16, S. 249). Da sich die entsprechend vorgenommenen Modifikationen jedoch in den beiden Teilstichproben nicht gleich auswirkten, wurden diese beiden Items bisher nicht entfernt. Beim zukünftigen Einsatz des StEB-Inventars sollten besonders diese Items erneut geprüft werden.

Der Faktor Entwicklung des Wissens erreicht mit .28 nur eine niedrige Varianz. Die Befragten unterscheiden sich zwar signifikant, aber nicht sehr stark in ihrer Ausprägung der Überzeugungen zur Entwicklung des erziehungswissenschaftlichen Wissens aus Studientexten.

Zentral für diese Arbeit ist die Fragestellung, ob sich epistemologische Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen aus Studientexten in der Faktorstruktur des StEB konform zum Theoriegefüge abbilden lassen (vgl. Kap 3, S. 135). Auf Basis der Ergebnisse von Studie 1 wurde diese Fragestellung bereits vorläufig beantwortet (siehe Kap. 6.5, S. 211). Die Faktorstruktur des StEB bildet in ihren inhaltlichen Dimensionen die epistemologischen Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen aus Studientexten ab und die resultierenden Dimensionen lassen sich im Hinblick auf das theoretische Kernkonstrukt inhaltlich sinnvoll beschreiben. Diese Antwort bleibt aufgrund der Stabilität der StEB-Faktoren auch für das finale 26-Item StEB Inventar gültig. Es lässt sich demnach festhalten, dass aus Studie 2 eine 26-Item StEB-Version resultiert, deren Faktorenstruktur inhaltlich der von Studie 1 entspricht. Das resultierende StEB Inventar verfügt folglich nach wie vor über die vier Dimensionen Überzeugungen zur *Absolutheit des Wissens*, *Einfachheit des Wissens*, *Multimodalität des Wissenserwerbs* und *Entwicklung des Wissens*. Diese vier Dimensionen entsprechen inhaltlich jenen, die aus der Studie 1 resultierten und sind folglich nach wie vor mit dem Theoriegefüge vereinbar (vgl. Kapitel 6.5, Abb. 6.4, S. 213). Aus Items zur *Quelle des Wissens* und *Rechtfertigung*

*des Wissens* bildete sich die StEB-Dimension *Absolutheit des Wissens*, die im übergeordneten theoretischen Bereich *nature of knowing* anzusiedeln ist. Items der Skalen *Sicherheit des Wissens* und *Struktur des Wissens* bilden gemeinsam die StEB-Dimension *Einfachheit des Wissens*, die in den Bereich *nature of knowledge* fällt (siehe Abb. 6.4, S. 213). Die neue StEB-Dimension *Multimodalität des Wissenserwerbs* setzt sich aus Items zusammen, die ursprünglich den Dimensionen *Quelle*, *Rechtfertigung* und *Sicherheit* zugeordnet waren. Für das finale StEB Inventar gilt demnach weiterhin die Skalenbeschreibung, die bereits in Kapitel 6.5 auf Seite 211 dargestellt ist.

Aus Tabelle 7.20 ist die Anzahl der Items je Faktor, die Anzahl der invertiert gepolten Items sowie die Anzahl der auf bestehenden Fragebögen basierenden Items (jedoch veränderten) je Skala zu entnehmen. Hier ist auch zu entnehmen, welche epistemologische Überzeugungsausprägung einer Zustimmung bzw. Ablehnung zu den Items der jeweiligen Subskala entspricht. Nach dem Entfernen ungeeigneter Items in Studie 2 verfügen nun alle vier Subskalen über durchweg gleichgerichtet formulierte Items. Über alle Items der Skalen *Absolutheit* und *Einfachheit* hinweg gilt zunächst (sofern diese nicht umgepolt werden), je höher die Zustimmung desto *weniger* sophistiziert ist die epistemologische Sichtweise auf Wissen. Bei den Skalen *Multimodalität* und *Entwicklung* verhält es sich anders herum, höhere Zustimmung stellt einheitlich für alle Items dieser Skalen eine sophistiziertere wissensbezogene Überzeugungsausprägung dar. Die Merkmalsausprägung, die inhaltlich einer Zustimmung zu den Items entspricht (rechter Pol), wurde für die Benennung der Subskalen verwendet.

TABELLE 7.20: Item-Skalen-Zuordnung des finalen 26-Item StEB Inventars

| Dimension | Item-anzahl | Itemanzahl invertiert | Itemanzahl abgewandelt | Ablehnung linker Pol | Zustimmung rechter Pol |
|---|---|---|---|---|---|
| Absolutheit des Wissens | 8 | 8 | 2 | relativ | absolut |
| Multimodalität des Wissenserwerbs | 7 | 0 | 1 | unimodal | multimodal |
| Einfachheit des Wissens | 6 | 6 | 0 | komplex | einfach |
| Entwicklung des Wissens | 5 | 0 | 2 | statisch | entwickelnd |
| Gesamt | 26 | 14 | 5 | | |

Der Wortlaut der 26 Items ist im Anhang B in den Tabellen B.16 und B.17 (S. 384, 385) einsehbar. Ebenfalls im Anhang B in Tabelle B.19 sind die fünf verbliebenen Items aufgeführt, die in ihrer Formulierung auf Items aus bestehenden Fragebögen basieren. Der Itemwortlaut dieser ursprünglichen Items ist dort ebenfalls einzusehen.

Zentral zur Beantwortung der Frage, ob sich epistemologische Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen aus Studientexten über das

StEB Inventar abbilden lassen, sind auch die psychometrischen Eigenschaften des resultierenden StEB Inventars. Im Folgenden werden einige dieser Eigenschaften im Rahmen der Item- und Skalenanalyse des finalen 26-Item StEB Inventars berichtet. Kapitel 8 (ab S. 273) bezieht sich insofern ebenfalls auf die Psychometrie des finalen StEB Inventars, als dass die Validierung der Testwertinterpretationen dargestellt wird.

## 7.6 Item- und Skalenanalyse

Die psychometrischen Eigenschaften der vier resultierenden Skalen und 26 verbleibenden Items werden in diesem Unterkapitel näher untersucht. Das umfasst nach der klassischen Testtheorie eine umfassende Item- und Skalenanalyse (Bühner, 2011). Zu Beginn wird geprüft, welches Messmodell gültig ist. Der Messmodellprüfung schließt sich die Skalenanalyse an, die eine Reliabilitätsanalyse sowie die Berechnung der Skalenhomogenitäten und Zusammenhänge zwischen den Faktoren umfasst. Darauf folgend wird diskutiert, wie die Testwerte des StEB zu bestimmen sind, bevor die Items im Hinblick auf ihre psychometrische Schwierigkeit und Trennschärfe untersucht werden.

### 7.6.1 Messmodellprüfung

Die Messmodelle der klassischen Testtheorie beschreiben die Art der Beziehung zwischen den manifesten und den latenten Variablen (siehe Kapitel 4.3.2, S. 160). Da die Reliabilitätsschätzung von den Annahmen des gültigen Messmodells abhängt, muss vor der Reliabilitätsanalyse ermittelt werden, von welchem Messmodell auszugehen ist. Die fünf bekanntesten Messmodelle sind das Modell paralleler, essentiell paralleler, tau-äquivalenter, essentiell tau-äquivalenter und tau-kongenerischer Messung (Bühner, 2011).

Wie in Kapitel 4 ab Seite 145 bereits beschrieben wurde, gibt jedes dieser fünf Messmodelle bestimmte Bedingungen vor. Die Bedingungen der Messmodelle betreffen die Intercepts (Itemmittelwerte = Itemschwierigkeit), Itemfehlervarianzen (Messgenauigkeit) und Ladungen der Items (Itemtrennschärfe). Es muss also identifiziert werden, wie die Beziehung zwischen den Items und dem entsprechenden zu messenden Merkmal am besten zu beschreiben ist. Die Modelle unterscheiden sich in der Strenge ihrer Annahmen. Das Modell paralleler Messung gibt die stärksten Annahmen vor. Nach der Reihe der Aufzählung werden die Modelle immer weniger restriktiv. Sind die Annahmen eines restriktiveren Modells erfüllt, können auch die weniger restriktiven Modelle, die weniger Bedingungen stellen, angenommen werden. In die andere Richtung gilt dies jedoch nicht.

Da die Messmodelle die Beziehung zwischen einem Faktor und den zugehörigen Items betrifft, sollte die Prüfung der Messmodelle zunächst je Faktor getrennt durchgeführt werden. Für die Prüfung müssen im einfaktoriellen Modell jeweils die entsprechenden Bedingungen der fünf Messmodellarten umgesetzt werden. Dann wird jeweils eine CFA berechnet und der Modell-Fit verglichen. Der Vergleich zeigt dann, unter den Bedingungen welchen Messmodells der beste Modell-Fit erreicht wurde. Dieses entsprechende Modell liefert dann die Basis für die Reliabilitätsanalyse.

Alle CFAs mit standardisierter Lösung (Varianz der latenten Variable auf 1 fixiert und Mittelwert auf 0 fixiert), die oben im Rahmen der Auswertung bereits durchgeführt wurden, entsprechen in ihren Bedingungen dem tau-kongenerischen Messmodell. Dieses Messmodell gibt keine Restriktionen für Intercepts, Ladungen oder Varianzen vor. Intercepts, Ladungen und Varianzen werden also nicht je auf einen Wert fixiert, sondern alle frei geschätzt. Inhaltlich bedeutet das bezüglich der Ladungen, dass die manifesten Variablen (Items) in

unterschiedlichem Ausmaß von der latenten Variable (dem zu messenden Merkmal) beeinflusst werden. Das heißt, dass die Items unterschiedlich trennscharf sind und demnach unterschiedlich hohe Ladungen resultieren. Die frei geschätzten Varianzen bedeuten inhaltlich für die Messung, dass davon ausgegangen wird, dass für die Items unterschiedlich hohe Messfehler vorliegen und demnach unterschiedliche hohe Messgenauigkeit ausweisen. Die Intercepts sind die durchschnittliche Itemantwort in der Stichprobe und deshalb ein guter Schätzer für die durchschnittliche Itemsschwierigkeit. Werden sie nicht gleichgesetzt, sondern frei geschätzt, heißt das inhaltlich für die Messung, dass die Items des Tests unterschiedlich hohe Schwierigkeiten aufweisen. Um den Bedingungen einer essentiell tauäquivalenten Messung zu genügen, müsste ein Modell guten Modell-Fit erreichen, bei dem die Ladungen aller Items gleichgesetzt wurden. Das würde dann bedeuten, dass anzunehmen ist, dass alle Items in gleich hohem Ausmaß von der latenten Variable beeinflusst werden und demnach alle über die gleiche Trennschärfe verfügen. Die Bedingungen der tauäquivalenten Messung bestehen in der Gleichsetzung der Ladungen und Intercepts. Alle Items verfügen über eine einheitliche Trennschärfe und die gleiche Itemschwierigkeit. Bei der essentiell parallelen Messung werden die Ladungen gleichgesetzt und auch die Fehlervarianzen auf einen Wert fixiert. Das parallele Messmodell ist das restriktivste. Hier werden sowohl die Ladungen ($\lambda_{i1}$), Fehlervarianzen ($\delta_i$) als auch Intercepts ($\lambda_{i0}$) jeweils gleichgesetzt. Es wird also davon ausgegangen, dass die Items alle gleich trennscharf, messgenau und schwierig sind.

Da die Messmodellprüfung je Faktor getrennt berechnet werden muss und zudem noch jeweils für die Gesamtstichprobe und die beiden Teilstichproben durchgeführt wurde, entspricht das insgesamt 60 berechneten CFAs. Die entsprechenden Modell-Fits sind im Anhang H in den drei Tabellen H.1, H.2 und H.3 ab Seite 438 einsehbar. Die Bedingungen des parallelen und tau-äquivalenten Messmodells führen für alle Stichproben zu deutlich schlechteren Modell-Fits als die der übrigen Messmodelle. Diese beiden Messmodelle unterscheiden sich von den übrigen Modellen dadurch, dass eine Fixierung der Intercepts vorgenommen wird. Das heißt, dass bei diesen Modellen für alle Items eines Faktors gleich hohe Itemschwierigkeiten (= mittlere Itembeantwortung über alle Befragten) angenommen werden. Aus diesen Ergebnissen lässt sich demnach ableiten, dass für alle Items aller vier Faktoren des resultierende StEB Inventars unterschiedlich hohe Itemschwierigkeiten vorliegen.

Das Modell mit dem besten Modell-Fit ist jeweils hervorgehoben. Hier zeigt sich, dass sowohl für die Gesamtstichprobe als auch für die beiden Teilstichproben für drei der vier Faktoren von einem tau-kongenerischen Messmodell ausgegangen werden kann. Bei den Faktoren *Absolutheit*, *Einfachheit* und *Multimodalität* führt das Weglassen von Restriktionen für Intercepts, Ladungen und Varianzen also zu einer besseren Modellpassung. Wenn unterschiedliche hohe Schwierigkeit, Trennschärfe und Messgenaugigkeit für die einzelnen Items zugelassen werden, und diese nicht gleichgesetzt werden, wird ein besserer Modell-Fit erreicht.

Bei dem Faktor *Entwicklung* führt dagegen die Bedingung des essentiell tau-äquivalenten Messmodells, die in der Gleichsetzung der Itemladungen bestehen, für alle Stichproben zu

einem besseren Modell-Fit. Es kann also davon ausgegangen werden, dass die Items des Faktors Entwicklung alle ähnlich trennscharf sind. Betrachtet man die Ladungen der standardisierten einfaktoriellen Lösung (siehe Anhang D, Abb. D.12, D.13 und D.14, S. 398), dann fällt auf, dass diese alle ähnlich hoch sind. Dementsprechend ergibt sich auch keine Verschlechterung des Modells, wenn sie auf den gleichen Wert fixiert werden. Für drei der vier Faktoren des StEB Inventars kann also das tau-kongenerische Messmodell angenommen werden. Für den Faktor *Entwicklung* kann das strengere essentiell tau-äquivalente Modell angenommen werden. Da die Annahme eines restriktiven Modells immer auch die Annahme der weniger restriktiven Modelle umfasst, kann für das StEB Inventar als Ganzes das tau-kongenerische Messmodell angenommen werden. Eine Überprüfung des finalen 26-Item vierfaktoriellen Gesamtmodells der Gesamtstichprobe zeigt[2], dass das unrestriktive tau-kongenerische Messmodell tatsächlich zur besten Modellpassung führt (siehe Tabelle 7.21). Dem folgt nach der klassischen Testtheorie, dass aus $X_i = \tau_i + \epsilon_i$, welches für jedes Item und jede zufällige Person als gültig angenommen wird, aufgrund der Annahme des tau-kongenerischen Modell $X_i = \lambda_{i0} + \lambda_{i1} \cdot \xi + \epsilon_1$ wird, da $\tau_i = \lambda_{i0} + \lambda_{i1} \cdot \xi$.

TABELLE 7.21: Messmodellprüfung des finalen vierfaktoriellen 26-Item Modells ($N$ = 915)

| Messmodelle | $\chi^2$ | *df* | $\chi^2/df$ | *p* | CFI | RMSEA | CI (90%) | CFit | SRMR |
|---|---|---|---|---|---|---|---|---|---|
| parall. | 4782.690 | 368 | 13 | .000 | .02 | .115 | .112; .117 | .000 | .181 |
| ess. parall. | 1463.278 | 343 | 4.27 | .000 | .75 | .060 | .057; .063 | .000 | .090 |
| $\tau$-äquiv. | 4999.619 | 343 | 14.58 | .000 | .16 | .109 | .107; .112 | .000 | .174 |
| ess.-$\tau$-äquiv. | 1027.538 | 318 | 3.23 | .000 | .84 | .049 | .046; .055 | .618* | .077 |
| **$\tau$-kong.** | 780.837 | 293 | 2.66 | .000 | .89 | .043 | .039; .046 | 1* | .052 |

[2]Genau genommen können Messmodelle nur falsifiziert und nicht wirklich bestätigt werden.

### 7.6.2 Skalenanalyse

Werden Reliabilitätskoeffizienten auf Basis der klassischen Testtheorie ermittelt, die Annahmen zum Messfehler und wahren Wert einer Messung aufstellt, handelt es sich immer um pauschale Genauigkeitsschätzungen der Testwerte (Döring & Bortz, 2016, S. 443). Nach der klassischen Testtheorie ist die Reliabilität von Messungen als Varianzverhältnis definiert. „Es handelt sich um den Anteil der Varianz der wahren Werte an der Varianz der beobachteten Werte." (Bühner, 2011, S. 147). Da die Varianzen für die Population aus der Stichprobe geschätzt werden müssen, handelt es sich demnach um eine Schätzung der Reliabilität (ebd.). Aufgrund der Tatsache, dass das einer Messung zugrundeliegende Messmodell, d.h. das spezifische Verhältnis zwischen erfasstem Merkmal und Items, auch bestimmte Bedingungen für die Messgenauigkeit der Items vorgibt, ist die adäquate Methode der Reliabilitätsschätzung vom Messmodell abhängig (Bühner, 2011).

Die Split-Half-, die Paralleltest- und die Retest-Methode eignen sich bei Modellen paralleler und essentiell paralleler Messung zur Bestimmung der Reliabilität[3] (vgl. Bühner, 2011, S. 153 ff.). Die Ermittlung der internen Konsistenz einer Skala mittels Cronbachs $\alpha$ (Cronbach, 1951) stellt die Standardmethode zur Schätzung der Skalenreliabilität dar. Cronbachs $\alpha$, dessen Berechnung auf Itemvarianzen und Itemkovarianzen beruht, stellt eine genaue Schätzung der Reliabilität dar, wenn mindestens das essentiell tau-äquivalente Modell gilt (Bühner, 2011, S. 158). Gilt, wie im vorliegenden Fall, ein schwächeres, tau-kongenerisches Modell, liefert Cronbachs $\alpha$ aber wenigstens eine Mindestschätzung der Reliabilität.

Im Falle der Reliabilitätsschätzung für die vier finalen Subskalen des StEB Inventars kann die Reliabilität demnach auch über Cronbachs $\alpha$ bestimmt werden, eine genauere Reliabilitätsschätzung liefert allerdings der $\omega$-Koeffizent (McDonald, 1999; Raykov, 2001a, 2001b), der sich explizit für tau-kongenerische Messungen eignet (vgl. Brown, 2015). In Tabelle 7.22 auf Seite 267 sind für die vier Subskalen des finalen 26-Item StEB Inventars einmal für die Gesamtstichprobe von $N$ = 915 sowie für die beiden Teilstichproben S1 und S2 jeweils der $\alpha$- und $\omega$-Koeffizent aufgeführt. Es zeigt sich, dass die Reliabilitätsschätzung $\omega$ für die Subskalen *Einfachheit des Wissens* und *Multimodalität des Wissenserwerbs* zu leicht höheren Ergebnissen führt. Die Reliabilitätsschätzungen für die Subskalen sind zwischen den Stichproben etwas unterschiedlich. Für die Gesamtstichprobe liegen die Reliabilitäten zwischen .70 bis .82.

Normalerweise werden Reliabilitäten von unter .80 als niedrig, von .80 bis .90 als mittel und ab .90 als hoch bewertet (Bühner, 2011, S. 81). Für die existierenden Skalen zur Erfassung epistemologischer Überzeugungen sind Reliabilitäten zwischen .70 und .80 jedoch sehr üblich, viele der verfügbaren Tests erreichen gar Reliabilitäten von $\alpha < .70$. Dies geht aus den Tabellen 2.13, 2.14, 2.15 und 2.16 in Kapitel 2.9.2 (ab S. 114) hervor. In Anbetracht dieser bisher erreichten Reliabilitätswerte für Skalen zur Erfassung epistemologischer Überzeugungen werden in Anlehnung an Nunnally (1978) bereits Reliabilitätswerte von .70 als ausreichend hoch erachtet (vgl. auch Klopp, 2014; Pintrich, 2002).

[3] Da die Modelle paralleler und essentiell paralleler Messung strenger sind als das tau-äquivalente Modell, kann in diesem Fall selbstverständlich auch Cronbachs $\alpha$ ermittelt werden.

TABELLE 7.22: Skalenreliabilität der finalen Subskalen des StEB

| Reliabilitätsmaße | Absolutheit | Einfachheit | Multimodalität | Entwicklung |
|---|---|---|---|---|
| GS ($N$ = 915) | | | | |
| $\alpha$ | .82 | .71 | .70 | .70 |
| $\omega 3$ | .82 | .73 | .70 | .70 |
| S1 ($N$ = 457) | | | | |
| $\alpha$ | .82 | .71 | .72 | .68 |
| $\omega 3$ | .82 | .74 | .71 | .68 |
| S2 ($N$ = 458) | | | | |
| $\alpha$ | .82 | .71 | .68 | .72 |
| $\omega 3$ | .82 | .72 | .70 | .72 |

In Tabelle 7.23 sind die *Mean-Inter-Item-Korrelationen* (*MIC*) als Maß für die Homogenitäten (Bühner, 2011, S. 243) der resultierenden StEB-Subskalen sowohl für die Gesamtstichprobe als auch die beiden Teilstichproben S1 und S2 dargestellt. Mit Werten von .25 bis .36. liegen die Skalenhomogenitäten des StEB alle innerhalb des von Briggs und Cheek (1986, S. 115) geforderten Akzeptanzbereichs von .20 bis .40.

TABELLE 7.23: Skalenhomogenität der finalen StEB-Subskalen getrennt nach Stichproben

| $MIC$ | Absolutheit | Einfachheit | Multimodalität | Entwicklung |
|---|---|---|---|---|
| *MIC* (GS $N$ = 915) | .36 | .30 | .25 | .32 |
| *MIC* (S1 $N$ = 457) | .36 | .30 | .27 | .29 |
| *MIC* (S2 $N$ = 458) | .36 | .30 | .27 | .29 |

### 7.6.3 Bestimmung der Testwerte des StEB

Im Rahmen eines Tests werden den Personen, die an dem Test teilnehmen, entsprechend ihrer Ausprägung des zu testenden Merkmals individuelle Testwerte zugeordnet. Diese Testwerte sollen demnach die Merkmalsausprägung widerspiegeln. Als Testwerte dienen häufig die über die Items einer Skala berechneten Summen- oder Mittelwerte. Teilweise werden die Items vor der Berechnung der Summen- oder Mittelwerte auch gewichtet. Eine andere Möglichkeit zur Berechnung der Testwerte stellen Faktorwerte dar. Von der Gültigkeit des Messmodells ist jedoch abhängig, wie die Testwerte sinnvollerweise berechnet werden müssen.

Die Gültigkeit des tau-kongenerischen Modells impliziert im vorliegenden Fall, dass die Items über unterschiedliche Trennschärfe (ungleiche Ladungen), Schwierigkeit (ungleiche Intercepts) und Messgenauigkeit (ungleiche Itemfehlervarianzen) verfügen. Dies würde zunächst bedeuten, dass Faktorwerte sich als Testwerte eignen würden, denn sie geben die Ausprägung einer Person auf einem Faktor an (Nunnally & Bernstein, 1994) und werden anhand der Faktorladungen der Items auf dem entsprechenden Faktor gewichtet und sind z-standardisiert ($M$ = 0, $SD$ = 1) (Bühner, 2011, S. 340). Sie berücksichtigen also die unterschiedlich hohen Ladungen der Items auf dem jeweiligen Faktor, indem hoch ladende Items, die das Zielkonstrukt besser vorhersagen, stärker in den Testwert eingehen, als niedriger ladende Items. Standardisierte Faktorwerte nehmen normalerweise Werte zwischen -3 und 3 an[4]. Hohe positive Faktorwerte stehen für eine hohe Ausprägung einer Person auf einem Faktor, d.h. in dem entsprechenden Zielkonstrukt, während hohe negative Werte für eine geringe Ausprägung stehen (Bühner, 2011, S. 340).

Die Bestimmung der Testwerte als Faktorwerte bringt allerdings bedeutende Nachteile für den Einsatz eines Tests mit sich, da diese aufgrund der Berechnung über die Ladungsgewichte stark stichprobenabhängig sind (Russell, 2002). Die Schätzung der Ladungen ist bei großen und für die Zielgruppe repräsentativen Stichproben genau und führt zu einer angemessenen Gewichtung der Items bei der Berechnung der Faktorwerte. Wird das StEB Inventar später eingesetzt, dann ist allerdings davon auszugehen, dass es nicht nur an großen und für die Population repräsentativen Stichproben eingesetzt werden wird, sondern auch an Stichproben von um die 30 bis 40 Personen. Bei so einer kleinen Stichprobe würden die Faktorladungen jedoch nicht mehr genau geschätzt werden (Russell, 2002), weshalb davon auszugehen ist, dass die entstandene Gewichtung der Items für diese kleine Stichprobe nicht angemessen wäre und sich von der Gewichtung in der hier vorliegenden großen Stichprobe unterscheiden würde. Die Verwendung von Faktorwerten ist demnach bei kleineren Stichproben problematisch (Bühner, 2011, S. 341) und für den späteren Einsatz unpraktikabel. Aus diesem Grund werden zur Bestimmung der Testwerte des StEB Inventars keine Faktorwerte gewählt, obwohl dies im Hinblick auf die Messmodellprüfung konsequent wäre. Stattdessen werden die Testwerte gebildet, indem für jede Dimension getrennt jeweils der Mittelwert über die entsprechenden Itemantworten einer Person berechnet wird.

[4]Faktorwerte können jedoch theoretisch von plus bis minus unendlich reichen (Bühner, 2011, S. 340).

Um sicherzustellen, dass diese Festlegung bezüglich der Testwerte die Ergebnisse der Testentwicklung nicht verfälscht, werden die Grundannahmen 5 bis 8, die der Validierung der Testwertinterpretationen dienen, sowohl für Faktorwerte als auch für Skalenmittelwerte vorgenommen und die resultierenden Ergebnisse verglichen (siehe Kapitel 8.4, S. 294).

### 7.6.4 Itemanalyse

Wie bereits im Zusammenhang mit der Itemanalyse der ersten Stichprobe (siehe Abschnitt 6.4.1, ab S. 206) dargestellt, dient bei Ratingskalen häufig der über alle Personen berechnete Itemmittelwert einer Stichprobe als Index für die psychometrische Schwierigkeit bzw. Leichtigkeit eines Items (Bühner, 2011, S. 219 f.). Da die vier Subskalen der resulierenden 26-Item StEB-Version aus inhaltlich gleichgerichteten Items bestehen, müssen die Items vor der Berechnung des Schwierigkeitsindex nicht umkodiert werden. Die Itemmittelwerte für die Gesamtstichprobe wurden bereits im Zusammenhang mit der deskriptiven Auswertung (siehe Tab. 7.8, S. 226) berichtet und sollen an dieser Stelle deshalb nicht wiederholt werden. Die Varianz bzw. Standardabweichung zeigt, wie stark die Itemantworten der Individuen vom Mittelwert abweichen (siehe Tab. 7.8, S. 226). Das Item mit dem höchsten Mittelwert ($M$ = 4.04 und $SD$ = 0.88) ist rw_27 („Es ist mir wichtig einen pädagogischen Studientext nicht nur zu lesen, sondern selbst zu durchdenken.“). Der für eine fünfstufige Skala hohe Mittelwert zeigt, dass viele Befragte dieser Aussage des Items zustimmen. Das umgekehrt formulierte Item sw_15_u („In pädagogischen Studientexten werden Dinge oft unnötig verkompliziert.“) wird dagegen mit einem Mittelwert von 2.00 ($SD$ = 0.97) eher abgelehnt.

Nach Moosbrugger (2012, S. 76) wurde außerdem der Schwierigkeitsindex $P_i$ berechnet, der bereits in Abschnitt 6.4.1 vorgestellt wurde. Zu berücksichtigen ist im vorliegenden Fall, dass zwei der Skalen ausschließlich aus umgekehrt formulierten Items bestehen. Bei den Skalen Einfachheit und Absolutheit drückt eine hohe Zustimmung aus, dass Wissen als absoluter bzw. als einfacher betrachtet wird. Die um die minimal erreichten Punkte korrigierten Schwierigkeitsindizes je Item für die Gesamtstichprobe sowie die beiden Teilstichproben sind Tabelle 7.24 auf Seite 270 zu entnehmen. Item rw_27 („Es ist mir wichtig einen pädagogischen Studientext nicht nur zu lesen, sondern selbst zu durchdenken.“) ist in der Gesamtstichprobe ($P_i$ = 75.93) und den ersten Teilstichproben ($P_i$ = 76.70) das psychometrisch einfachste Item. In der zweiten Teilstichprobe erreicht dagegen Item sw_15_u („In pädagogischen Studientexten werden Dinge oft unnötig verkompliziert.“) mit 75.22 den höchsten $P_i$-Wert (niedrigste Schwierigkeit). Item qw_25_u („Ich zweifle nicht an, was in pädagogischen Studientexten steht.“) verfügt einheitlich über die höchste psychometrische Schwierigkeit mit $P_i$-Werten von 40.21 bis 40.88.

TABELLE 7.24: Itemschwierigkeit getrennt nach Stichproben

| Items | Itemschwierigkeitsindex $P_i$ | | |
|---|---|---|---|
| | GS ($N$ = 915) | S1 ($N$ = 457) | S2 ($N$ = 458) |
| Absolutheit | | | |
| rw_33_u | 54.89 | 53.17 | 56.6 |
| rw_29_u | 51.83 | 51.86 | 51.8 |
| qw_25_u | 40.55 | 40.21 | 40.88 |
| qw_21_u | 51.45 | 50.66 | 52.24 |
| qw_26_u | 48.93 | 48.30 | 49.56 |
| qw_13_u | 43.17 | 42.12 | 44.21 |
| gw_4_u | 63.22 | 62.58 | 63.86 |
| rw_18_u | 53.44 | 52.74 | 54.15 |
| Einfachheit | | | |
| sw_15_u | 74.75 | 74.29 | 75.22 |
| gw_13_u | 53.22 | 52.63 | 53.82 |
| sw_2_u | 73.17 | 72.37 | 73.96 |
| sw_17_u | 68.36 | 67.51 | 69.21 |
| rw_36_u | 50.36 | 48.85 | 51.86 |
| gw_3_u | 65.66 | 65.75 | 65.56 |
| Multimodalität | | | |
| qw_11 | 71.15 | 70.51 | 71.78 |
| gw_9 | 70.82 | 69.97 | 71.67 |
| rw_27 | 75.93 | 76.70 | 75.16 |
| rw_12 | 66.09 | 66.25 | 65.94 |
| rw_11 | 73.91 | 73.96 | 73.85 |
| sw_22 | 73.83 | 73.85 | 73.8 |
| gw_16 | 57.6 | 57.06 | 58.13 |
| Entwicklung | | | |
| ew_13 | 71.97 | 71.06 | 72.87 |
| ew_12 | 68.17 | 67.89 | 68.45 |
| ew_3 | 60.03 | 59.74 | 60.32 |
| ew_8 | 55.25 | 56.07 | 54.42 |
| ew_10 | 63.14 | 63.57 | 62.72 |

Die korrigierte Item-Skala-Korrelation – die Itemtrennschärfe $r_{ix}$ – gibt an, wie gut das einzelne Item das Gesamtergebnis eines Tests repräsentiert (siehe Kap. 6.4.2 (ab S. 209). Items mit hoher Trennschärfe trennen erfolgreich zwischen Gruppen mit niedriger und hoher Merkmalsausprägung (Döring & Bortz, 2016, S. 476 ff.). In Tabelle 7.25 wird die part-whole korrigierte Itemtrennschärfe dargestellt. Diese wurde je Item in Bezug auf die betreffende Subskala berechnet und sowohl für die Gesamtstichprobe als auch für die beiden Teilstichproben ermittelt. Die Items erreichen in beiden Teilstichproben jeweils die gleichen Trennschärfewerte. Die Skala *Entwickung* besteht fast durchgehend aus Items mit mittlerer Trennschärfe ($r_{ix}$ > .30) (Bühner, 2011, S. 81), während in der Subskala *Absolutheit* Items mit hoher ($r_{ix}$ > .50) Trennschärfe vorherrschen. Subskala *Einfachheit* enthält zwei Items mit hoher ($r_{ix}$ > .50) und vier mit mittlerer (.30 bis .50) Trennschärfe. Die Items gw_3_u und gw_16 verfügen mit .31 und .29 über die geringsten Trennschärfewerte der verbleibenden 26 Items. Item gw_16 erreicht in der Gesamtstichprobe mit .29 nur einen niedrigen Trennschärfewert, während es jedoch in den beiden Teilstichproben mit jeweils .33 im Bereich mittlerer Trenschärfe liegt.

TABELLE 7.25: Korrigierte Itemtrennschärfe $r_{ix}$ der 26 verbleibenden StEB-Items je Stichprobe

| $r_{ix}$ | | | Item | Subskala |
|---|---|---|---|---|
| GS ($N$ = 915) | S1 ($N$ = 457) | S2 ($N$ = 458) | | |
| .50 | .51 | .51 | rw_33_u | Absolutheit |
| .59 | .59 | .59 | rw_29_u | Absolutheit |
| .61 | .59 | .59 | qw_25_u | Absolutheit |
| .52 | .49 | .49 | qw_21_u | Absolutheit |
| .61 | .58 | .58 | qw_26_u | Absolutheit |
| .53 | .53 | .53 | qw_13_u | Absolutheit |
| .46 | .49 | .49 | gw_4_u | Absolutheit |
| .48 | .50 | .50 | rw_18_u | Absolutheit |
| .40 | .39 | .39 | sw_15_u | Einfachheit |
| .50 | .50 | .50 | gw_13_u | Einfachheit |
| .54 | .56 | .56 | sw_2_u | Einfachheit |
| .51 | .53 | .53 | sw_17_u | Einfachheit |
| .41 | .40 | .40 | rw_36_u | Einfachheit |
| .33 | .31 | .31 | gw_3_u | Einfachheit |
| .56 | .52 | .52 | qw_11 | Multimodalität |
| .37 | .44 | .44 | gw_9 | Multimodalität |
| .43 | .48 | .48 | rw_27 | Multimodalität |
| .43 | .46 | .46 | rw_12 | Multimodalität |
| .39 | .39 | .39 | rw_11 | Multimodalität |
| .40 | .39 | .39 | sw_22 | Multimodalität |
| .29 | .33 | .33 | gw_16 | Multimodalität |
| .44 | .39 | .39 | ew_13 | Entwicklung |
| .41 | .40 | .40 | ew_12 | Entwicklung |
| .51 | .49 | .49 | ew_3 | Entwicklung |
| .47 | .44 | .44 | ew_8 | Entwicklung |
| .42 | .44 | .44 | ew_10 | Entwicklung |

# 8 Validierung der Testwertinterpretationen des StEB

Diese Arbeit stützt sich auf das aktuelle Verständnis des Validitätsbegriffs nach den *Standards for Educational and Psychological Testing* der AERA et al. (2014), das im Rahmen der Beschreibung des methodischen Hintergrunds dieser Arbeit in Kapitel 4.1 (ab S. 147) bereits vorgestellt wurde. Nach diesem argumentationsbasierten Ansatz gliedert sich die Validierung in vier Teilschritte, die insgesamt das Validitätsargument bilden. In einem ersten Schritt wird spezifiziert, wie die Testergebnisse zu interpretieren sind. Daraus werden empirisch prüfbare Grundannahmen abgeleitet und geprüft. Die Ergebnisse dieser Prüfung der Grundannahmen stellen empirische Evidenzen für oder gegen diese Grundannahmen zur Verfügung, die wiederum entweder die Gültigkeit der spezifizierten Testwertinterpretationen stützen oder aber gegen deren Gültigkeit sprechen. In einer zusammenfassenden Beurteilung werden die Ergebnisse dieses Validierungsprozesses abschließend festgehalten. So entwickelt sich in einer stringenten Argumentation die Validierung der angestrebten Testwertinterpretationen (vgl. Hartig et al., im Druck).

Die Spezifikation der Testwertinterpretationen und die Ableitung der prüfbaren Grundannahmen wurde in Kapitel 3 (ab S. 135) bereits vorgenommen. In diesem Kapitel wird die Prüfung der acht Grundannahmen dargestellt. Alle empirischen Evidenzen werden zusammengetragen und die Gültigkeit der Grundannahmen und somit der Testwertinterpretationen wird zusammenfassend bewertet.

Diese Testentwicklung zielt darauf ab, dass das StEB Inventar sich dazu eignet, die bei Lehramtsstudierenden aktuell vorherrschenden Ausprägungen in den Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten mehrdimensional zu ermitteln (siehe Kap. 3, ab S. 135). Der Einsatz des StEB Inventars soll darüber hinaus sowohl einen intraindividuellen als auch einen interindividuellen Vergleich der mehrdimensionalen Ausprägungen epistemologischer Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten ermöglichen. Das StEB Inventar soll folglich dazu genutzt werden können, Aussagen darüber zu treffen, ob ein Lehramtsstudierender im Vergleich zu einem anderen über weniger sophistizierte (niedrige Merkmalsausprägung) oder aber sophistiziertere (höhere Merkmalsausprägung) epistemologische Überzeugung zu erziehungswissenschaftlichem Wissen aus Studientexten (in den verschiedenen Dimensionen) verfügt. Über einen wiederholten Einsatz des StEB bei denselben Lehramtsstudierenden soll es außerdem möglich sein, individuelle Veränderungen in den Ausprägungen epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten (zu den verschiedenen

Dimensionen) zu identifizieren. Diese verschiedenen intendierten Interpretationen der Testergebnisse und geplanten Arten der Testnutzung des StEB Inventars müssen im Rahmen dieser Arbeit folglich validiert werden. Die Testwertinterpetationen, die es zu validieren gilt, sind:

**Testwertinterpretation 1 (erklärender Schluss):** Die mit dem StEB Inventar erfassten Testwerte sind Indikatoren für die Dimensionen epistemologischer Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten. Die Höhe der Ausprägung spiegelt sich in den Testwerten wider (vgl. Kap. 4.2, ab S. 150).

**Testwertinterpretation 2 (bewertender Schluss):** Die Testwerte des StEB Inventars können dazu eingesetzt werden, Personen mit geringerer Merkmalsausprägung, also weniger sophistizierten epistemologischen Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten, von Personen mit höherer Merkmalsausprägung, also sophistizierteren epistemologischen Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten, zu unterscheiden (vgl. Kap. 4.2, ab S. 150).

Als Datengrundlage für die folgende Prüfung der sieben Grundanannahmen dient der in Kapitel 7.4 (S. 221) bereits beschriebene, bereinigte Datensatz der zweiten Studie mit $N$ = 915 Lehramtsstudierenden.

## 8.1 Testinhalt und interne Struktur des StEB

Für die Prüfung der Grundannahmen 1 und 2 dienen die interne Struktur und der Testinhalt des StEB Inventars als Evidenzquellen. Die Teststruktur wurde über die konfirmatorische Faktorenanalyse in Studie 2 untersucht (siehe Abschnitt 7.5.2 ab S. 239). Nach dem alten Validitätsverständnis würden sich die hier dargestellten Aspekte hauptsächlich auf die faktorielle Validität im Bereich der Konstruktvalidierung beziehen.

**Grundannahme 1:** Die Antworten im StEB (Testwerte) lassen sich kausal auf verschiedene Überzeugungsdimensionen zurückführen.

**Grundannahme 2:** Das System resultierender Überzeugungsdimensionen zum erziehungswissenschaftlichem Wissen aus Studientexten ist mit der theoretischen Konzeptionalisierung nach Hofer und Pintrich (1997) vereinbar.

Bezüglich Grundannahme 1 konnte über die signifikanten unstandardisierten Ladungen der Items auf den Faktoren und die signifikanten Varianzen der vier Faktoren des finalen StEB CFA-Modells gezeigt werden (siehe Parameterschätzungen im Anhang E, Tab. E.7 bis E.9, ab S. 406), dass das Antwortverhalten auf den Items durch die jeweilige latente Variable bedingt wird. Das bedeutet inhaltlich, dass die Itembeantwortung einer Person kausal auf die vorliegende Ausprägung der Überzeugung zu erziehungswissenschaftlichem Wissen aus Studientexten (der spezifischen Dimension) zurückzuführen ist (Grundannahme 1). Die geschätze Itemladung gibt an, um wie viel einer Likertstufe (z.B. um .88 einer Likertstufe) sich der Wert dieses Items durchschnittlich ändert, wenn sich die Merkmalsausprägung in der entsprechenden Überzeugungsdimension um einen Punkt auf der Likertskala ändert. Das zeigt demnach, dass die erfassten Testwerte des StEB Inventars Indikatoren für die Dimensionen epistemologischer Überzeugungen sind (Grundannahme 1, Testwertinterpretation 1).

Die Item-Faktor-Zuordnung hat sich darüber hinaus auch als stabil herausgestellt. Die exploratorische Faktorenanalyse in Studie 2 hat die in der CFA spezifizierte Faktorstruktur bestätigt. Mittels einer Stichprobenteilung konnten Modellanpassungen kreuzvalidiert werden, um so die Stabilität der finalen StEB-Lösung sicherzustellen.

Die signifikanten geschätzten Varianzen der vier Faktoren des finalen StEB-Modells (siehe Anhang E, Tab. E.11, S. 409) bedeuten inhaltlich, dass sich die Versuchspersonen in ihren Merkmalsausprägungen – in ihren Überzeugungen zur Absolutheit, Einfachheit und Entwicklung des Wissens sowie Multimodalität des Wissenserwerbs – signifikant unterscheiden. Dies stellt eine wichtige Voraussetzung dafür da, dass die erfassten Testwerte des StEB Inventars Indikatoren für die Dimensionen epistemologischer Überzeugungen sein können (Grundannahme 1, Testwertinterpretation 1), denn wenn der Test nicht zwischen Personen mit höheren und niedrigeren Merkmalsausprägungen trennen könnte, wäre er ganz grundlegend ungeeignet, das Zielkonstrukt zu erfassen. Dabei handelt es sich auch um eine Voraussetzung dafür, dass sich das StEB Inventar dafür eignet, intraindividuelle und interindividuelle Unterschiede zu identifizieren (Testwertinterpretation 2).

Auch die Itemtrennschärfen, die Bestimmtheitsmaße ($R^2$) und Varianzen der Items stellen empirische Evidenzen zur Verfügung, mittels derer sich die beiden Grundannahmen prüfen lassen. Im Rahmen der Itemanalyse des finalen StEB wurden die part-whole korrigierten Itemtrennschärfen ($r_{ix}$) ermittelt (siehe Kap. 7.6.4, S. 269). Die Itemtrennschärfe $r_{ix}$ gibt an, wie gut das einzelne Item das Gesamtergebnis eines Tests repräsentiert. Items mit hoher Trennschärfe trennen erfolgreich zwischen Gruppen mit niedriger und hoher Merkmalsausprägung (Döring & Bortz, 2016, S. 476 ff.). In der Subskala *Absolutheit* herrschen Items mit hoher ($r_{ix}$ > .50) Trennschärfe vor. Die Subskala *Einfachheit* enthält zwei Items mit hoher ($r_{ix}$ > .50) und 4 mit mittlerer (.30 bis .50) Trennschärfe. Die Skala *Entwickung* besteht fast durchgehend aus Items mit mittlerer Trennschärfe ($r_{ix}$ > .30) (Bühner, 2011, S. 81). Kritisch zu bewerten ist Item gw_16 (*Multimodalität*), das mit $r_{ix}$ = .29 für die Gesamtstichprobe nur eine niedrige Trennschärfe erreicht, demnach den Gesamttest nicht so gut repräsentiert und weniger erfolgreich zwischen Gruppen mit niedriger und hoher Merkmalsausprägung unterscheidet.

Die Bestimmtheitsmaße ($R^2$) der Items des finalen StEB-Modells geben Aufschluss über den Varianzanteil, der durch die entsprechende latente Variable erklärt wird (vgl. Bühner, 2011, S. 453). Diese Werte sind als Kommunalitäten, das heißt als Mindestschätzungen der Itemreliabilität zu interpretieren (ebd.). Inhaltlich bedeutet das, dass Items mit einem hohen $R^2$ durch ihre jeweilige latente Variable besser vorhergesagt werden, als Items mit niedrigen $R^2$. Als hohe Itemkommunalitäten gelten Werte von .60 - .80, als mittlere Werte von .40 - .60 und niedrige Kommunalitäten entsprechen Werten von .20 - .40 (siehe Bühner, 2011, S. 345). Die Bestimmtheitsmaße ($R^2$) der Items des finalen StEB (siehe Anhang E, Tab. E.9, S. 407) bewegen sich zum Großteil zwischen .20 - .40. Fünf Items weisen Werte über .40 auf. Wie bereits diskutiert (siehe 7.5.3, S. 258), verfügen die Items gw_3_u und gw_16 mit $R^2$ = .16 bzw. $R^2$ = .12 jedoch über sehr niedrige Werte. Liegen hauptsächlich niedrige Kommunalitäten vor, kann das nach Bühner (2011, S. 453) ein Hinweis darauf sein, dass es sich um einen heterogenen Faktor handelt. Im vorliegenden Fall handelt es sich bei den Faktoren *Multimodalität*, *Einfachheit* und *Absolutheit* um heterogene Faktoren, da sie sich aus Itemsubsets zusammensetzen.

Mit den niedrigen Bestimmtheitsmaßen ($R^2$) der Items gw_3_u und gw_16 gehen hohe standardisierte Itemvarianzen ($1 - R^2$) einher (siehe Anhang E, Tab. E.11, S. 409). Sie verfügen also über hohe itemspezifische Varianz, die sich nicht durch die betreffende Ausprägung in der Überzeugung erklären lässt und demnach Fehlervarianz darstellt.

Die aufgeführten empirischen Evidenzen zeigen an, dass Grundannahme 1 anzunehmen ist und die Antworten im StEB demnach auf die Ausprägungen in den Überzeugungsdimensionen zurückzuführen sind. Dies stützt die angenommenen Testwertinterpretationen. Die Evidenzen zeigen allerdings auch, dass die Items gw_3_u und gw_16 sich weniger gut eignen, die betreffende Überzeugungsausprägung (gw_3_u, Einfachheit; gw_16, Multimodalität) zu erfassen. Sie sollten beim Einsatz des StEB kritisch untersucht werden.

Zur Prüfung von Grundannahme 2 ist zunächst festzustellen, dass die Itemformulierung zu den angenommenen Kernbereichen der epistemologischen Überzeugungen *Quelle des Wissens*, *Sicherheit des Wissens*, *Struktur des Wissens* und *Rechtfertigung des Wissens* (Hofer & Pintrich, 1997) sowie zum Bereich *Entwicklung des Wissens* (Conley et al., 2004) erfolgte. Der Testinhalt wurde direkt nach der Itementwicklung geprüft, indem der Itempool zehn Expertinnen und Experten vorgelegt wurde, welche die Items in Bezug auf ihre Eignung zur intendierten Skala einschätzten (vgl. Kap. 5.3, S. 179). So konnte in Studie 1 ein Testentwurf des StEB mit optimierter Itemauswahl eingesetzt werden. In Studie 1 konnte anhand der EFA bereits gezeigt werden, dass die grundsätzliche Struktur des Instruments, d.h. die Zuordung der Items zu den Subkalen, mit dem Theoriegefüge vereinbar ist, auch wenn nicht die fünf ursprünglich intendierten Subskalen identifiziert wurden (vgl. Kap. 6.5, ab S. 211). Die vier resultierenden Überzeugungsdimensionen lassen sich einerseits im Hinblick auf das Zielkonstrukt als *Absolutheit des Wissens*, *Einfachheit des Wissens*, *Multimodalität des Wissenserwerbs* und *Entwicklung des Wissens* sinnvoll interpretieren und beschreiben. Andererseits fügen sich diese StEB-Skalen in die theoretisch angenommene, übergeordnete Struktur der *nature of knowing* und *nature of knowledge* ein (Hofer & Pintrich, 1997). Wie auf Basis der Ergebnisse der zweiten Studie in Kapitel 7.5.2 (ab S.235) außerdem gezeigt werden konnte, ließ sich diese Teststruktur anhand einer zweiten, größeren Stichprobe replizieren. Sowohl die Zuordnung der Items zu den Faktoren[1], als auch die Beschreibung der inhaltlichen Dimensionen blieben in Studie 2 erhalten (vgl. Kap. 7.5.3, ab S. 258). Dies zeigt, dass die resultierende Faktorstruktur stabil ist. Grundannahme 2 kann demnach als bestätigt gelten, was wiederum die Annahme stärkt, dass die Testwertinterpretationen gültig sind.

[1]Mit Ausnahme des Items ew_10, das im Zuge der Modellanpassung in Studie 2 wieder seinem ursprünglich intendierten Faktor *Entwicklung des Wissens* zugeordnet wurde und nicht bei dem Faktor verblieb, dem es in Studie 1 zugeordnet wurde.

## 8.2 Zusammenhänge zwischen den StEB-Dimensionen

Zur Prüfung von Grundannahme 3 werden die Beziehungen zwischen den vier finalen Skalen des StEB Inventars untersucht. Nach dem alten Validitätsverständnis, würde sich dieser Aspekt auf die Prüfung der divergenten Validität beziehen (vgl. Bortz & Döring, 2006, S. 203).

**Grundannahme 3:** Zwischen den Überzeugungsdimensionen des StEB liegen moderate Zusammenhänge vor, da die Überzeugungsdimensionen verschiedene Teilaspekte des übergeordneten Konstrukts *epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten* abbilden.

In der CFA, die für die Validierungsstichprobe durchgeführt wurde (Studie 2), wurden die Kovarianzen zwischen den vier Faktoren der finalen 26-Item StEB-Version bestimmt[2]. Diese Kovarianzen, die zugehörigen Standardfehler (*SE*) sowie *z*- und *p*-Werte, die Aufschluss über die Beziehung zwischen den Subskalen untereinander geben, sind in Tabelle 8.1 dargestellt.

TABELLE 8.1: Kovarianzen zwischen den Faktoren nach dem Ergebnis der CFA der 26-Item StEB-Version

| Subskalen | geschätzte *cov* | *SE* | *z* | *p* |
|---|---|---|---|---|
| Absolutheit mit | | | | |
| Einfachheit | .235 | .031 | 7.561 | .000 |
| Multimodalität | -.157 | .025 | -6.205 | .000 |
| Entwicklung | -.094 | .022 | -4.373 | .000 |
| Einfachheit mit | | | | |
| Multimodalität | -.176 | .027 | -6.438 | .000 |
| Entwicklung | -.004 | .018 | -0.221 | .825 |
| Multimodalität mit | | | | |
| Enwicklung | .136 | .018 | 7.661 | .000 |

Bei der Interpretation ist zu berücksichtigen, dass die Faktoren *Absolutheit* und *Einfachheit* beide aus Items bestehen, die so formuliert sind, dass eine hohe Zustimmung *weniger* sophistizierte epistemologische Überzeugungen zum erziehungswissenschaftlichen Wissen ausdrückt, während eine starke Ablehnung hier für größere Sophistiziertheit spricht. Bei den anderen beiden Faktoren, *Multimodalität* und *Entwicklung*, drückt eine hohe Zustimmung dagegen hohe Sophistiziertheit der wissensbezogenen Überzeugungen aus, eine starke Ablehnung dagegen weniger sophistizierte Überzeugungen. Am größten ist der signifikant positive Zusammenhang ($p < .001$) zwischen den Faktoren *Einfachheit* und *Absolutheit*. Der Faktor *Absolutheit* hängt signifikant negativ mit den Faktoren *Multimodalität* und *Entwicklung* zusammen ($p < .001$), die jedoch ihrerseits untereinander einen signifikant positiven Zusammenhang aufweisen ($p < .001$). Die Faktoren *Multimodalität* und *Einfachheit* sind

[2]Sie wurden im Rahmen der CFA Ergebnisse in Abschnitt 7.5.2 jedoch noch nicht dargestellt.

signifikant negativ assoziiert ($p < .001$). Nur zwischen den beiden Subskalen *Einfachheit* und *Entwicklung* findet sich kein nennenswerter ($p > .05$) Zusammenhang.

Zur Prüfung der Grundannahme 3, die postuliert, dass bei Gültigkeit der Testwertinterpretationen Zusammenhänge zwischen den Subskalen des StEB auftreten müssten, ist zunächst zu prüfen, ob die bestehenden Zusammenhänge sich inhaltlich sinnvoll, d.h. konstruktkonform zueinander verhalten. Nach dem alten Validitätsverständnis wäre dieser Aspekt in den Bereich der Konstruktvalidierung bzw. auch der Inhaltsvalidierung einzuordnen gewesen (vgl. Testgütekriterien S. 148).

Wie aus Tabelle 8.1 (S. 278) zu entnehmen ist, besteht ein signifikanter positiver Zusammenhang ($p < .001$) zwischen den Dimensionen *Überzeugungen zur Absolutheit des Wissens* und *Einfachheit des Wissens*. Die Überzeugung, dass erziehungswissenschaftliche Studientexte eine verlässliche Quelle für erziehungswissenschaftliches Wissen darstellen, deren Inhalte nicht anzuzweifeln sind, geht mit der Überzeugung einher, dass dieses Wissen aus sicheren, diskreten Fakten besteht, die es zu wissen gilt. Dieser Zusammenhang ist einerseits in seiner Richtung (positiv) sinnvoll, da eine hohe Zustimmung in diesen beiden Dimensionen jeweils für unsophistizierte, naive wissensbezogene Überzeugungen spricht. Auf der anderen Seite ist dieser Zusammenhang auch inhaltlich sinnvoll, da in Bezug auf das Konstrukt nachvollziehbar ist, dass ein Lehramtsstudierender, der glaubt, dass erziehungswissenschaftliches Wissen im Kern aus sicheren, diskreten Fakten besteht, gleichzeitig auch der Überzeugung ist, dass diese Fakten verlässlich und nicht anzuzweifeln sind.

Sinnvoll ist auch der signifikante negative Zusammenhang ($p < .001$) zwischen den gegensätzlich gerichteten Dimensionen *Überzeugungen zur Absolutheit des Wissens* und *Multimodalität des Wissenserwerbs*. Die Überzeugung, dass erziehungswissenschaftliche Studientexte eine verlässliche Quelle für erziehungswissenschaftliches Wissen darstellen, deren Inhalte nicht anzuzweifeln sind, ist negativ assoziiert mit der Überzeugung, dass erziehungswissenschaftliches Wissen für das eigene Verstehen einer vielfältigen, elaborierten Auseinandersetzung mit verschiedenen Wissensquellen, Forschungsmeinungen und Sichtweisen bedarf. Auch dies ist in Bezug auf das Konstrukt nachvollziehbar, denn wieso sollte eine Person, die glaubt, dass Zweifel an den erziehungswissenschaftlichen Inhalten unangebracht sind, sich mit anderen Meinungen und Sichtweisen auseinandersetzen wollen? Schließlich ist diese Person ja der Überzeugung, schon die „richtige“ Sichtweise gefunden zu haben. Das eigene Verstehen und die Auseinandersetzung mit verschiedenen Ansichten ist nicht so wichtig, wenn man der Meinung ist, dass die Inhalte einfach so hingenommen werden sollten, weil sie verlässlich sind. Dieser Person scheint für den eigenen Wissenserwerb das einfache Lesen erziehungswissenschaftlicher Texte als angemessene Auseinandersetzung.

Die Überzeugung zur Multimodalität des Wissenserwerbs hängt darüber hinaus negativ mit der Überzeugung zur Einfachheit des Wissens zusammen. Folglich ist eine Lehramtsstudierende, die der Überzeugung ist, dass erziehungswissenschaftliches Wissen aus Studientexten aus sicheren, diskreten Fakten besteht, die es zu wissen gilt (einfache Wissensstruktur), eher der Ansicht, dass das einfache Lesen eines Studientextes als Auseinandersetzung reicht (im Gegensatz zur vielfältigen, elaborierten Auseinandersetzung). Auch diese Beziehung der beiden Dimensionen ist in sich schlüssig.

Ebenso schlüssig ist der negative Zusammenhang zwischen der Dimension *Absolutheit des Wissens* und der *Entwicklung des Wissens*. Ein Studierender, der der Überzeugung ist, dass erziehungswissenschaftliche Studientexte eine verlässliche Quelle für erziehungswissenschaftliches Wissen darstellen, deren Inhalte nicht anzuzweifeln sind, geht eher davon aus, dass erziehungswissenschaftliches Wissen über die Zeit unverändert bleibt und eine erneute Auseinandersetzung mit den Inhalten unnötig ist.

Schließlich liegt noch ein positiver Zusammenhang zwischen der Entwicklung des Wissens und der Multimodalität des Wissenserwerbs vor. Auch dieser Zusammenhang fügt sich inhaltlich stringent in das Bild des Gesamtkonstrukts ein. Lehramtsstudierende, die der Überzeugung sind, dass sich das erziehungswissenschaftliche Wissen über die Zeit verändert und einer erneuten Auseinandersetzung bedarf, sind eher auch der Überzeugung, dass für den erziehungswissenschaftlichen Wissenserwerb eine vielfältige, elaborierte Auseinandersetzung mit verschiedenen Wissensquellen, Forschungsmeinungen und Sichtweisen nötig ist.

Zwischen der Dimension *Überzeugungen zur Entwicklung des Wissens* und *Einfachheit des Wissens* besteht kein Zusammenhang. Offensichtlich ist es für Lehramtsstudierende, die davon ausgehen, dass sich erziehungswissenschaftliches Wissen weiterententwickelt, nicht gleichzeitig wahrscheinlicher anzunehmen, dass dieses Wissen auch *komplex strukturiert* ist (unsichere, untereinander verbundene Konzepte) statt *einfach strukturiert* (sichere, diskrete zu wissende Fakten). Auf den ersten Blick könnte dieser Nullzusammenhang widersprüchlich erscheinen. Dieser scheinbare Widerspruch klärt sich jedoch über den Rückgriff auf die Theorie auf. Die Dimension *Einfachheit des Wissens* betrifft nach Hofer und Pintrich (1997) die Natur des Wissens (*nature of knowledge*) und nicht die Natur des Wissensprozesses (*nature of knowing*). Der Bereich *Natur des Wissens* betrifft die Frage, wie Wissen in sich aufgebaut ist, im Sinne von einfach oder komplex[3]. Fokus der Dimension *Einfachheit des Wissens* ist die Überzeugung darüber, wie Wissen aufgebaut ist, nicht aber, wie man sich Wissen aneignet. Schließlich ist es durchaus denkbar, dass eine Person annimmt, dass erziehungswissenschaftliches Wissen aus sicheren, diskreten Fakten besteht, diese Fakten jedoch nach einiger Zeit durch neue Fakten abgelöst werden, die ab dem Zeitpunkt ihrer Einführung gelten und dann für eine gewisse Zeit sicher sind. Es ist dagegen auch denkbar, dass eine Person, die annimmt, dass Wissen aus komplexen, verbundenen Konzepten besteht, gleichzeitig entweder der Überzeugung sein kann, dass dieses komplexe Zusammenspiel aus Konzepten gleich bleibt oder aber, dass sich dieses komplexe Geflecht über die Zeit verändert.

Wie aus der Prüfung der Zusammenhänge hervorgeht, fügen sich diese inhaltlich stringent in das Bild des Gesamtkonstrukts ein. Die Grundannahme 3, dass Zusammenhänge zwischen den Subskalen des StEB bestehen, da diese verschiedene Teilaspekte eines übergeordneten Konstrukts messen, lässt sich demnach bestätigen und stützt die Gültigkeit der Testwertinterpretationen.

[3] Andere Autoren nennen vergleichbare Skalen *Struktur des Wissens* oder aber *Komplexität des Wissens*.

## 8.3 Beziehungen zwischen den StEB- und LIST-R-Testwerten

Um Grundannahme 4 zu prüfen, werden die Beziehungen zwischen den Dimensionen epistemologischer Überzeugungen des StEB Inventars mit dem berichteten Lernstrategieeinsatz untersucht. Nach dem alten Validitätsverständnis würde sich diese Prüfung auf die konvergente Validität beziehen.

**Grundannahme 4:** Die epistemologischen Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten weisen moderate Zusammenhänge mit der berichteten präferierten Lernstrategieauswahl beim Lernen mit erziehungswissenschaftlichen Studientexten auf, wobei sophistiziertere Ausprägungen epistemologischer Überzeugungen (in den verschiedenen Dimensionen) mit tiefenorientierten Lernstrategien (Elaborationsstrategien, metakognitive Lernstrategien und Organisationsstrategien) einhergehen und weniger sophistizierte Ausprägungen epistemologischer Überzeugungen mit der oberflächenorientierten Wiederholungsstrategie einhergehen.

In diesem Zusammenhang werden die empirischen Beziehungen zwischen den Testkennwerten des StEB Inventars und den Testkennwerten des LIST-R (Inventar zur Erfassung von Lernstrategien im Studium) (Wild, 2011) mit der theoretisch angenommenen Zusammenhangsstruktur verglichen, die aus den empirischen Ergebnissen früherer Studien abgeleitet wurden. Die auf Basis der Theorie angenommene und in zahlreichen empirischen Studien bestätigte Zusammenhangsstruktur ist in Kapitel 2.10 in Abbildung 2.18 (S. 130) dargestellt. Die Prüfung der empirischen Zusammenhangsstruktur des StEB und LIST-R soll zunächst über Strukturgleichungsmodelle und im Anschluss korrelativ vorgenommen werden (vgl. Hartig et al., 2012, S. 158). Zunächst wird im folgenden jedoch knapp auf die Datenbereinigung und deskriptive Analyse sowie die Reliabilitäts- und Itemanalyse der LIST-R-Skalen eingegangen.

Wie bereits in Kapitel 7.2 (S. 219) beschrieben, wurden die Lehramtsstudierenden im Rahmen der zweiten, größeren Erhebung zu ihrem Lernstrategieeinsatz in Bezug auf ihre erziehungswissenschaftlichen Studienanteile befragt. Dazu wurde eine gekürzte und revidierte Fassung des LIST Inventars von Wild und Schiefele (1994) zur Erfassung von Lernstrategien im Studium namens LIST-R von Wild (2011) eingesetzt. Der LIST-R deckt mit seinen 53 Items die drei übergeordneten Bereiche *kognitive Lernstrategien*, *metakognitive Lernstrategien* und *ressourcenbezogene Strategien* ab. Der Bereich der kognitiven Lernstrategien umfasst die drei Skalen *Organisieren* (AL.IO), *Wiederholen* (AL.IWE) und *Elaborieren* (AL.IVP) mit je sieben Items zu diesen drei Strategiearten. Die metakognitiven Lernstrategien (LO.MP, LO.MQ) werden über eine Skala mit insgesamt zehn Items erfasst. Drei dieser Items beziehen sich auf die aktive Planung und sieben auf die Regulation des eigenen Lernverhaltens. Der Bereich der ressourcenbezogenen Strategien besteht aus fünf Subskalen mit je vier Items zu *Lernzeit* (RN.LZ), *Literatur* (RN.LN), *Online-Ressourcen* (RN.OQ) und *kooperativem Lernen* (RN.GL) sowie sechs Items zur *Lernumgebung* (RN.LU) (siehe Kapitel 7.2, Tab. 7.1).

### 8.3.1 Bereinigung der Daten und deskriptive Analyse des LIST-R

Als Grundlage für die Analyse der Zusammenhänge zwischen den StEB- und LIST-R-Skalen dient der bereinigte Datensatz der zweiten Stichprobe mit $N = 915$ Lehramtsstudierenden. Von den 915 Datensätzen mussten 16 aufgrund vollständig fehlender Angaben bei den LIST-R-Items entfernt werden. Weitere acht Datensätze wurden aufgrund eines übermäßig hohen Anteils fehlender Werte bei den LIST-R Items – 21% bis 60% fehlend – ausgeschlossen. Demnach besteht die Stichprobe der folgenden Analysen aus $N = 891$ befragten Lehramtsstudierenden (553 weiblich, 333 männlich, 5 keine Angabe) im durchschnittlichen Alter von 22 Jahren ($M = 21.98$, $SD = 3.27$). Das mittlere Fachsemester ist $M = 3.13$ ($SD = 2.73$) und das mittlere Hochschulsemester $M = 4.57$ ($SD = 3.92$). Die Ergebnisse der deskriptiven Analyse dieser Stichprobe ($N = 891$) für die StEB-Items sind im Anhang I in Tabelle I.1 auf Seite 441 dargestellt.

Die folgenden Tabellen 8.2 und 8.3 auf den Seiten 283 und 284 zeigen die Ergebnisse der deskriptiven Analyse für die LIST-R-Items. Der Wortlaut der LIST-R-Items ist im Anhang C in den Tabellen C.1 und C.2 (ab S. 390) zu sehen. Für den Aufbau des LIST-R und die Beschreibung der Skalen sei auf Kapitel 7.2 (S. 219) verwiesen. Die durchschnittliche angegebene Lernzeit für die erziehungswissenschaftlichen Studienanteile in Stunden pro Woche während des Semesters (LZSemester) beträgt $M = 9.81$ ($SD = 9.21$), in der vorlesungsfreien Zeit (LZFerien) dagegen $M = 7.26$ ($SD = 11.46$). Auffällig ist der jeweils sehr große Range von 70 für LZSemester bzw. 88 für LZFerien. Aufgrund der sehr rechtssteilen Verteilung beschreiben allerdings die Mediane von $Md = 7$ für LZSemester und $Md = 3.5$ für LZFerien die angegebene Lernzeit pro Woche besser in dieser Stichprobe. Die teilweise sehr hohen Stundenangaben von beispielsweise 40 bis 88 Stunden Lernzeit pro Woche geben deutliche Hinweise darauf, dass diese beiden Items von vielen Befragten entweder nicht ernsthaft beantwortet wurden, auf alle Studieninhalte und nicht nur die erziehungswissenschaftlichen Studienanteile bezogen wurden oder die Befragten ihre eigene Lernzeit nur sehr schwer realistisch einschätzen können.

TABELLE 8.2: Deskriptive Analyse der LIST-R-Items für die Konstruktvalidierung (bereinigte Stichprobe, $N = 891$) – Teil 1

| Items | *n* | *M* | *SD* | *Md* | min | max | range | skew | kurtosis | *SE* |
|---|---|---|---|---|---|---|---|---|---|---|
| LZSemester | 867 | 9.81 | 9.21 | 7 | 0 | 70 | 70 | 2.31 | 7.78 | .313 |
| LZFerien | 859 | 7.26 | 11.46 | 3.5 | 0 | 88 | 88 | 4.07 | 22.6 | .391 |
| AL.IO1 | 890 | 2.75 | 1.22 | 3 | 1 | 5 | 4 | 0.11 | -0.91 | .041 |
| AL.IO2 | 889 | 4.33 | 0.91 | 5 | 1 | 5 | 4 | -1.52 | 2.14 | .031 |
| AL.IO3 | 888 | 3.42 | 1.19 | 4 | 1 | 5 | 4 | -0.33 | -0.85 | .04 |
| AL.IO4 | 890 | 4.11 | 0.92 | 4 | 1 | 5 | 4 | -1.03 | 0.8 | .031 |
| AL.IO5 | 891 | 3.94 | 1.02 | 4 | 1 | 5 | 4 | -0.82 | 0.13 | .034 |
| AL.IO6 | 890 | 3.06 | 1.21 | 3 | 1 | 5 | 4 | -0.06 | -0.93 | .041 |
| AL.IO7 | 890 | 3.08 | 1.33 | 3 | 1 | 5 | 4 | -0.12 | -1.16 | .045 |
| AL.IWE1 | 889 | 4.34 | 0.79 | 4 | 1 | 5 | 4 | -1.31 | 2.06 | .026 |
| AL.IWE2 | 891 | 4.25 | 0.9 | 4 | 1 | 5 | 4 | -1.32 | 1.58 | .03 |
| AL.IWE3 | 891 | 3.97 | 1 | 4 | 1 | 5 | 4 | -0.86 | 0.22 | .034 |
| AL.IWE4 | 891 | 2.69 | 1.31 | 3 | 1 | 5 | 4 | 0.3 | -1.04 | .044 |
| AL.IWE5 | 891 | 3.71 | 1.1 | 4 | 1 | 5 | 4 | -0.63 | -0.32 | .037 |
| AL.IWE6 | 891 | 4.39 | 0.74 | 5 | 1 | 5 | 4 | -1.36 | 2.49 | .025 |
| AL.IWE7 | 890 | 3.08 | 1.23 | 3 | 1 | 5 | 4 | -0.05 | -1.02 | .041 |
| AL.IVP1 | 891 | 3.44 | 0.99 | 4 | 1 | 5 | 4 | -0.39 | -0.25 | .033 |
| AL.IVP2 | 891 | 3.57 | 0.9 | 4 | 1 | 5 | 4 | -0.34 | -0.2 | .03 |
| AL.IVP3 | 891 | 4.05 | 0.8 | 4 | 1 | 5 | 4 | -0.65 | 0.29 | .027 |
| AL.IVP4 | 890 | 3.64 | 0.98 | 4 | 1 | 5 | 4 | -0.39 | -0.38 | .033 |
| AL.IVP5 | 890 | 2.38 | 1 | 2 | 1 | 5 | 4 | 0.51 | -0.11 | .034 |
| AL.IVP6 | 890 | 2.72 | 1.08 | 3 | 1 | 5 | 4 | 0.26 | -0.6 | .036 |
| AL.IVP7 | 891 | 2.98 | 1.03 | 3 | 1 | 5 | 4 | 0 | -0.42 | .034 |
| LO.MP1 | 891 | 3.55 | 1.14 | 4 | 1 | 5 | 4 | -0.63 | -0.37 | .038 |
| LO.MP2 | 891 | 3.67 | 1.02 | 4 | 1 | 5 | 4 | -0.6 | -0.12 | .034 |
| LO.MP3 | 891 | 3.88 | 1.01 | 4 | 1 | 5 | 4 | -0.87 | 0.28 | .034 |
| LO.MQ1 | 890 | 3.31 | 1.05 | 3 | 1 | 5 | 4 | -0.25 | -0.56 | .035 |
| LO.MQ2 | 891 | 3.91 | 0.97 | 4 | 1 | 5 | 4 | -0.73 | 0.07 | .032 |
| LO.MQ3 | 891 | 2.87 | 1.07 | 3 | 1 | 5 | 4 | 0.09 | -0.6 | .036 |
| LO.MQ4 | 891 | 4 | 0.8 | 4 | 1 | 5 | 4 | -0.76 | 0.96 | .027 |
| LO.MQ5 | 890 | 3.75 | 0.95 | 4 | 1 | 5 | 4 | -0.68 | 0.14 | .032 |
| LO.MQ6 | 890 | 4.3 | 0.73 | 4 | 1 | 5 | 4 | -1.14 | 2.28 | .025 |
| LO.MQ7 | 891 | 4.28 | 0.74 | 4 | 1 | 5 | 4 | -0.88 | 0.81 | .025 |

TABELLE 8.3: Deskriptive Analyse der LIST-R-Items für die Konstruktvalidierung (bereinigte Stichprobe, $N$ = 891) – Teil 2

| Items | *n* | *M* | *SD* | *Md* | min | max | range | skew | kurtosis | *SE* |
|---|---|---|---|---|---|---|---|---|---|---|
| RN.LZ1 | 891 | 2.84 | 1.14 | 3 | 1 | 5 | 4 | 0 | -0.84 | .038 |
| RN.LZ2 | 891 | 3.18 | 1.19 | 3 | 1 | 5 | 4 | -0.26 | -0.91 | .04 |
| RN.LZ3 | 891 | 2.44 | 1.19 | 2 | 1 | 5 | 4 | 0.5 | -0.7 | .04 |
| RN.LZ4 | 891 | 2.73 | 1.23 | 3 | 1 | 5 | 4 | 0.09 | -1.1 | .041 |
| RN.LU1 | 891 | 4.42 | 0.71 | 5 | 1 | 5 | 4 | -1.3 | 2.34 | .024 |
| RN.LU2 | 891 | 4.35 | 0.71 | 4 | 2 | 5 | 3 | -0.85 | 0.3 | .024 |
| RN.LU3 | 890 | 3.67 | 1.08 | 4 | 1 | 5 | 4 | -0.46 | -0.58 | .036 |
| RN.LU4 | 891 | 4.18 | 0.83 | 4 | 1 | 5 | 4 | -0.92 | 0.69 | .028 |
| RN.LU5 | 891 | 3.94 | 1.06 | 4 | 1 | 5 | 4 | -0.81 | -0.07 | .035 |
| RN.LU6 | 889 | 4.25 | 0.85 | 4 | 1 | 5 | 4 | -1.09 | 0.93 | .029 |
| RN.LN1 | 891 | 3.09 | 1.16 | 3 | 1 | 5 | 4 | -0.11 | -0.82 | .039 |
| RN.LN2 | 890 | 3.69 | 1.1 | 4 | 1 | 5 | 4 | -0.73 | -0.18 | .037 |
| RN.LN3 | 890 | 3.25 | 1.14 | 3 | 1 | 5 | 4 | -0.23 | -0.74 | .038 |
| RN.LN4 | 890 | 3.06 | 1.14 | 3 | 1 | 5 | 4 | -0.02 | -0.7 | .038 |
| RN.OQ1 | 891 | 3.97 | 1.03 | 4 | 1 | 5 | 4 | -0.91 | 0.23 | .034 |
| RN.OQ2 | 891 | 3.93 | 1.02 | 4 | 1 | 5 | 4 | -0.8 | 0.07 | .034 |
| RN.OQ3 | 891 | 4.43 | 0.73 | 5 | 1 | 5 | 4 | -1.46 | 2.92 | .025 |
| RN.OQ4 | 891 | 2.92 | 1.24 | 3 | 1 | 5 | 4 | 0.08 | -0.95 | .042 |
| RN.GL1 | 891 | 2.71 | 1.19 | 3 | 1 | 5 | 4 | 0.13 | -0.84 | .04 |
| RN.GL2 | 891 | 3.07 | 1.13 | 3 | 1 | 5 | 4 | -0.15 | -0.79 | .038 |
| RN.GL3 | 891 | 3.74 | 1.02 | 4 | 1 | 5 | 4 | -0.66 | 0.05 | .034 |
| RN.GL4 | 890 | 3.81 | 1.01 | 4 | 1 | 5 | 4 | -0.7 | 0.08 | .034 |

### 8.3.2 Reliabilitätsanalyse der LIST-R-Skalen

Nach der Bereinigung der Daten wurden zunächst die internen Konsistenzen der neun Subskalen des LIST-R bestimmt. Diese sind in Tabelle 8.4 dargestellt und werden hier den Skalenreliabilitäten gegenübergestellt, die Wild (2011) in seinem Kurzmanual zum LIST-R berichtet. Sie beziehen sich bei Wild (2011) auf eine Stichprobe von $N$ = 436 Studierenden aus wirtschaftswissenschaftlichen Studiengängen.

In der vorliegenden Stichprobe von $N$ = 891 Lehramtsstudierenden erreichen die vier Subskalen *Elaborieren*, *Lernzeit*, *Literatur* und *Kooperatives Lernen* mit $\alpha > .80$ jeweils mittlere Reliabilitäten (Fisseni, 1997). Die übrigen sind mit $\alpha < .80$ als niedrig (Fisseni, 1997) einzustufen. Die interne Konsistenz der Skala *Online-Ressourcen* ist mit $\alpha = .66$ als inakzeptabel gering einzustufen. In Bezug auf diese Skala weichen die $\alpha$-Werte zwischen beiden Stichproben am stärksten ab. Wird Item 1 („Ich lerne an einem Platz, an dem ich mich wohl fühle.“) aus der Skala *Lernumgebung* entfernt, steigt Cronbachs $\alpha$ leicht an. Bei der Skala *Organisieren* betrifft dies die Items 1 und 7 („Ich fertige Tabellen, Diagramme oder Schaubilder an, um den Stoff besser strukturiert vorliegen zu haben.“, „Ich stelle wichtige Fachausdrücke und Definitionen in eigenen Listen zusammen.“). Auch hier verbessert sich die interne Konsistenz, wenn diese Items entfernt werden. Diese angepassten $\alpha$-Werte sind in Tabelle 8.4 in Klammern dargestellt. Für die Subskalen *Elaborieren* und *Literatur* wurden in der vorliegenden Stichprobe etwas höhere Reliabilitäten gefunden als Wild (2011) diese berichtet.

Die Skalen *Elaborieren*, *Lernzeit*, *Literatur* und *Kooperatives Lernen* sind mit Werten von $MIC > .40$ (siehe Tabelle 8.4) sehr homogen und liegen damit über dem von Briggs und Cheek (1986, S. 115) geforderten Akzeptanzbereich von .20 bis .40. Die übrigen LIST-R-Skalen bewegen sich in ihren Skalenhomogenitäten innerhalb des Akzeptanzbereichs.

TABELLE 8.4: Skalenhomogenität und interne Konsistenz der LIST-R-Skalen im Vergleich Wild (2011)

| Lernstrategien | $N$ = 891 | | $N$ = 436 (Wild, 2011) |
|---|---|---|---|
| | $\alpha$ | $MIC$ | $\alpha$ |
| Organisieren | .71 (.73) | .27 | .79 |
| Wiederholen | .74 (.75) | .29 | .74 |
| Elaborieren | .84 | .42 | .82 |
| Metakognitive Strategien | .74 | .23 | .81 |
| Lernzeit | .83 | .56 | .86 |
| Lernumgebung | .74 (.75) | .33 | .79 |
| Literatur | .88 | .64 | .86 |
| Online-Ressourcen | .66 | .36 | .78 |
| Kooperatives Lernen | .84 | .57 | .86 |

In Klammern sind die angepassten $\alpha$ bei Entfernen der Items: IO1, IO7 für *Organisieren*, IWE1, IWE2, IWE6 für *Wiederholen* und LU1 für *Lernumgebung*.

In Tabelle 8.5 werden die part-whole korrigierten Trennschärfen der Items des LIST-R dargestellt. Items mit mittlerer und hoher Trennschärfe dominieren die Skalen. Nur drei der 53 Items IO7, IWE6 und LU1 verfügen über eine geringe Trennschärfe ($r_{ix} < .30$). Die Items IO1 und IWE2 erreichen mit $r_{ix} = .31$ und $r_{ix} = .32$ nur knapp mittlere Trennschärfe ($.30 < r_{ix} < .50$, Bühner, 2011, S. 81). Bei diesen fünf genannten Items mit geringer oder relativ geringer Trennschärfe handelt es sich – neben einem weiteren Item – um diejenigen, die die interne Konsistenz senken (vgl. Tab. 8.4, S. 285). Ein Entfernen dieser Items aus den jeweiligen Skalen hat einen Anstieg von Cronbachs $\alpha$ zur Folge. Die Skalen *Elaborieren*, *Lernzeit*, *Literatur* und *Kooperatives Lernen* bestehen ausschließlich aus Items mit hoher Trennschärfe (siehe Tabelle 8.5). Dies sind gleichzeitig auch die Skalen mit sehr hoher Homogenität (vgl. Tab. 8.4, S. 285).

TABELLE 8.5: Part-whole korrigierte Itemtrennschärfe der LIST-R-Items ($N = 891$)

| Kognitive Lernstrategien | | Metakognitive Lernstrategien | | Ressourcenbez. Strategien | |
|---|---|---|---|---|---|
| Items je Skala | $r_{ix}$ | Items je Skala | $r_{ix}$ | Items je Skala | $r_{ix}$ |
| Organisieren | | Lernen organisieren | | Lernzeit | |
| IO1 | .31 | MP1 | .35 | LZ1 | .68 |
| IO2 | .50 | MP2 | .44 | LZ2 | .67 |
| IO3 | .53 | MP3 | .41 | LZ3 | .70 |
| IO4 | .39 | MQ1 | .39 | LZ4 | .59 |
| IO5 | .46 | MQ2 | .37 | Lernumgebung | |
| IO6 | .53 | MQ3 | .36 | LU1 | **.29** |
| IO7 | **.27** | MQ4 | .46 | LU2 | .52 |
| Wiederholen | | MQ5 | .42 | LU3 | .51 |
| IWE1 | .45 | MQ6 | .43 | LU4 | .59 |
| IWE2 | .32 | MQ7 | .46 | LU5 | .50 |
| IWE3 | .47 | | | LU6 | .51 |
| IWE4 | .55 | | | Literatur | |
| IWE5 | .59 | | | LN1 | .76 |
| IWE6 | .30 | | | LN2 | .67 |
| IWE7 | .53 | | | LN3 | .72 |
| Elaborieren | | | | LN4 | .79 |
| IVP1 | .61 | | | Online-Ressourcen | |
| IVP2 | .59 | | | OQ1 | .38 |
| IVP3 | .55 | | | OQ2 | .53 |
| IVP4 | .52 | | | OQ3 | .53 |
| IVP5 | .61 | | | OQ4 | .41 |
| IVP6 | .64 | | | Kooperatives Lernen | |
| IVP7 | .58 | | | GL1 | .62 |
| | | | | GL2 | .63 |
| | | | | GL3 | .74 |
| | | | | GL4 | .70 |

Hervorgehobene Werte: niedrige korrigierte Trennschärfe $r_{ix} < .30$ (Bühner, 2011, S. 81)

### 8.3.3 CFA des LIST-R

Vor der Durchführung der konfirmatorischen Faktorenanalyse der LIST-R-Items müssen zunächst die geltenden Voraussetzungen bezüglich Kollinearität und Verteilung für diese Analyse geprüft werden (siehe Abschnitt 7.5.1, S. 229 für eine genauere Beschreibung der Voraussetzungen).

Sehr hoch miteinander korrelierende Items von $r >| .85 |$ können im Rahmen der CFA zu Schätzproblemen führen. Aus diesem Grund sollte vor der Durchführung der CFA sichergestellt werden, dass sogenannte Kollinearität nicht vorliegt. Eine Inspektion der Itemkorrelationsmatrix zeigt, dass von Kollinearität auszugehen ist. Die Itemkorrelationen (Pearson) bewegen sich von minimal r = -.18 bis maximal r = .86. Als problematisch ist der sehr hohe Zusammenhang ($r >| .85 |$) zwischen dem Item 3 und 4 der Skala *Kooperatives Lernen* anzusehen. Um Schätzproblemen im Rahmen der CFA vorzubeugen, werden die beiden betroffenen Items sowie die beiden übrigen Items der Skala *Kooperatives Lernen* – zu der ohnehin keine Fragestellung oder Grundannahme im Rahmen dieser Arbeit vorliegen – von den weiteren Analysen ausgeschlossen. Multivariate Normalverteilung liegt für die 53 LIST-R-Items mit $p < .001$ nicht vor. Die LIST-R-Items liegen jedoch mit Werten für die Schiefe von -1.52 bis 0.51 und Exzess von -1.16 bis 2.92 deutlich innerhalb der von West et al. (1995) vorgeschlagenen Grenzen (Schiefe < 2, Exzess < 7).
Die ML-Methode kann somit trotz der Verletzung der multivariaten Normalverteilung angewandt werden. Es wird aufgrund des Likert-Formats der manifesten Variablen und der formalen Verletzung der Normalverteilungsannahme der robuste ML-Schätzer (MLR) gewählt. Bezüglich des Umgangs mit fehlenden Werten wird die FIML-Methode angewandt. Im Rahmen der ersten Modellspezifikation werden die reliabelsten Items je Subskala identifiziert, um diese als Referenzitems definieren zu können. Der Großteil der LIST-R-Items lädt eindeutig, mit einer geschätzten Ladung von $\lambda_{i1} > .40$, auf den jeweilig intendierten Faktoren. Die drei Items AL.IWE2 („Ich lese meine Aufzeichnungen mehrmals hintereinander durch.“), AL.IWE6 („Ich gehe in Gedanken noch einmal den gelernten Stoff durch, um zu sehen, ob ich alles Wesentliche behalten habe.“), RN.LU1 („Ich lerne an einem Platz, an dem ich mich wohlfühle.“) verfügen mit $\lambda_{i1} < .40$ über geringe Ladungen. Die geschätzten Ladungen aller Items sind signifikant ($p < .001$). Auch die Varianzen aller neun Faktoren sind signifikant. Das neunfaktorielle Gesamtmodell erreicht jedoch eine unzureichende Modellpassung von: $\chi^2(1289) = 4524.151$, $p < .001$, $\chi^2/df = 3.51$, RMSEA = .053 (CI90: .051;.055), CFit = .001 (RMSEA < .05), CFI = .78, SRMR = .065. Der CFI ist deutlich unterhalb der Grenze von .90 (vgl. Bewertungsrichtlinien für die Modellpassung Tab. 4.3, S. 165)

Eine Prüfung der einfaktoriellen Modelle für die 9 LIST-R-Faktoren ergibt darüber hinaus, dass sie alle unzureichende Modellpassung erzielen sowie korrelierende Itemfehlervarianzen und hohe geschätzte Kovarianzresiduen aufweisen und deshalb alle neun abgelehnt werden müssen. Die Modellpassungen der einfaktoriellen LIST-R-Modelle sind in Tabelle 8.6 auf Seite 288 einzusehen.

TABELLE 8.6: Modell-Fit der einfaktoriellen LIST-R-Modelle und des neunfaktoriellen Gesamtmodells

| Modelle | $\chi^2$ | *df* | $\chi^2/df$ | *p* | CFI | RMSEA | CI (90%) | CFit | SRMR |
|---|---|---|---|---|---|---|---|---|---|
| Kognitive Lernstrategien | | | | | | | | | |
| Organisieren | 357.741 | 14 | 25.55 | .000 | .732 | .166 | .152; .181 | .000 | .080 |
| Wiederholen | 182.070 | 14 | 13.01 | .000 | .832 | .116 | .103; .130 | .000 | **.064** |
| Elaborieren | 337.587 | 14 | 24.11 | .000 | .834 | .161 | .147; .175 | .000 | **.067** |
| Metakognitive Strategien | | | | | | | | | |
| Lernen organisieren | 503.441 | 35 | 14.38 | .000 | .687 | .123 | .114; .131 | .000 | .086 |
| Ressourcenbezogene Strategien | | | | | | | | | |
| Lernzeit | 16.077 | 2 | 8.04 | .000 | **.986** | .089 | .057; .125 | .024 | **.019** |
| Lernumgebung | 385.078 | 9 | 42.79 | .000 | .680 | .217 | .201; .233 | .000 | .105 |
| Literatur | 22.248 | 2 | 11.12 | .000 | **.985** | .107 | .075; .142 | .002 | **.017** |
| Online-Ressourcen | 18.993 | 2 | 9.50 | .000 | **.968** | .098 | .064; .136 | .012 | **.032** |
| Kooperatives Lernen | 189.584 | 2 | 94.79 | .000 | .886 | ..324 | .290; .361 | .000 | .089 |
| Gesamtmodell | 4524.151 | 1289 | 3.51 | .000 | .778 | **.053** | **.051; .055** | .001 | **.065** |

Hervorgehobene Werte befinden sich innerhalb der Cut-off Werte von Hu und Bentler (1999)

Eine Prüfung der Beziehungen der LIST-R-Subskalen und StEB-Subskalen mittels Strukturgleichungsmodellen mit dem Ziel der Prüfung von Grundannahme 4 ist aufgrund der fehlenden Modellpassung der LIST-R-Faktoren nicht möglich. Selbst wenn nur die vier in diesem Fall theoretisch relevanten LIST-R-Faktoren (Elaboration, Organisieren, Wiederholen und metakognitive Strategien) in einem Strukturmodell neben den StEB-Skalen spezifiziert werden würden (oder auch nur je ein LIST-R-Faktor mit einem einzigen StEB-Faktor), würde dies aufgrund der schlechten Modellpassung der LIST-R-Faktoren unweigerlich zu einer Ablehnung des Modells führen.

### 8.3.4 Analyse der Zusammenhänge zwischen StEB und LIST-R

Zur Prüfung der korrelativen Beziehungen zwischen dem berichteten Lernstrategieeinsatz und den epistemologischen Überzeugungen der Lehramtsstudierenden zu erziehungswissenschaftlichem Wissen, werden die Zusammenhänge zwischen den neun Dimensionen des LIST-R und den vier Dimensionen des StEB ermittelt. Hervorzuheben ist an dieser Stelle, dass die Studierenden beim Ausfüllen des LIST-R im Rahmen der Instruktion aufgefordert wurden, ihre Aussagen zum Lernstrategieeinsatz ausschließlich auf ihre erziehungswissenschaftlichen Studienanteile zu beziehen. Gleichzeitig ist auch das StEB Inventar ausschließlich auf den erziehungswissenschaftlichen Bereich des Lehramtsstudiums bezogen, und hier noch genauer fokussiert auf die Studientexte aus dem Bereich Erziehungswissenschaft. Diese Bereichsspezifität ist für die Interpretation der Zusammenhänge zwischen dem berichteten Lernstrategieeinsatz und den Überzeugungen zur Absolutheit, Einfachheit, Entwicklung sowie zur Multimodalität des Wissenserwerbs von zentraler Bedeutung.

Für den LIST-R werden die individuellen Merkmalsausprägungen berechnet, indem für jede Subskala getrennt die Summe über die Itemwerte dieser Skala berechnet wird (siehe Wild, 2011). Auch für das StEB Inventar müssen die individuellen Merkmalsausprägungen für jede Skala getrennt berechnet werden. Wie in Abschnitt 7.6.3 (S. 268) diskutiert, werden für das StEB Inventar als Testwerte die individuellen Skalenmittelwerte berechnet. Tabelle 8.7 (S. 290) sind die Pearson-Korrelationskoeffizienten zwischen den vier Skalenmittelwerten des StEB und den neun Skalensummenwerten des LIST-R zu entnehmen. Dass die Zusammenhänge signifikant werden, ist aufgrund der Stichprobengröße wenig verwunderlich. Allerdings interessiert im Rahmen dieser Betrachtung vielmehr die im Hinblick auf die Theorie inhaltlich sinnvolle Interpretierbarkeit der identifizierten Zusammenhänge. Wie sich zeigt, liegen signifikante Korrelationen zwischen allen vier Skalen des StEB und dem berichteten Einsatz von kognitiven Elaborationsstrategien vor. Die Überzeugungen zur Absolutheit und Einfachheit des Wissens korrelieren beide signifikant negativ mit dem berichteten Einsatz von Elaborationsstrategien. Das bedeutet, dass Lehramtsstudierende, die der (wenig sophistizierten) Überzeugung sind, dass erziehungswissenschaftliche Studientexte eine verlässliche Quelle des Wissens darstellen, deren Inhalte man ohne zu hinterfragen glauben sollte (Absolutheit), einen geringeren Einsatz von Elaborationsstrategien für ihr Lernen von erziehungswissenschaftlichen Studienanteilen berichten, als Studierende mit sophistizierten relativen Überzeugungen in dieser Dimension.

TABELLE 8.7: Korrelationskoeffizienten (Pearson) zwischen den Skalenmittelwerten des StEB und den Skalensummenwerten des LIST-R

| Subskalen | Absolutheit | Einfachheit | Multimodalität | Entwicklung |
|---|---|---|---|---|
| Kognitive LS | | | | |
| Organisieren | .09 (.07) | -.03 (-.06) | .19** (.18**) | .07 (.06) |
| Wiederholen | .17** (.20**) | .23** (.24**) | -.02 (-.07) | .08 (.04) |
| Elaborieren | -.25** | -.29** | .42** | .15** |
| Metakognitive LS | | | | |
| Lernen organisieren | .06 | 0 | .21** | .13** |
| Ressourcenbezogene LS | | | | |
| Lernzeit | .15** | .04 | .01 | -.03 |
| Lernumgebung | .08 | .03 (.04) | .09 (.08) | .06 (.04) |
| Literatur | -.05 | -.19** | .33** | .01 |
| Online-Ressourcen | .03 | .15** | .06 | .04 |
| Kooperatives Lernen | .02 | -.02 | .06 | .06 |

In Klammern sind die angepassten Korrelationskoeffizienten $r$ bei Entfernen der Items IO1, IO7 für *Organisieren*, IWE1, IWE2, IWE6 für *Wiederholen* und LU1 für *Lernumgebung*.
** $p < .01$

Die Überzeugung, erziehungswissenschaftliches Wissen aus Studientexten bestehe aus sicheren, diskreten zu wissenden Fakten, geht ebenfalls mit einem geringeren (berichteten) Einsatz von Elaborationsstrategien für das Lernen von erziehungswissenschaftlichen Studienanteilen einher. Zwischen den beiden StEB-Skalen Überzeugungen zur *Multimodalität des Wissenserwerbs* und *Entwicklung des Wissens* mit der Skala *Elaboration* besteht jeweils ein signifikanter positiver Zusammenhang. Lehramtsstudierende, die der Überzeugung sind, dass der erziehungswissenschaftliche Wissenserwerb einer vielfältigen, elaborierten Auseinandersetzung mit den Studientexten bedarf (im Gegensatz zu einfachem Durchlesen), berichten mehr Elaborationsstrategien beim Lernen ihrer erziehungswissenschaftlichen Studienanteile einzusetzen. Ebenso verhält es sich für Lehramtsstudierende, die der Überzeugung sind, dass erziehungswissenschaftliches Wissen sich weiterentwickelt und eine erneute Auseinandersetzung erfordert. Bezüglich der Dimension *Multimodalität des Wissenserwerbs* bestehen darüber hinaus auch signifikante positive Zusammenhänge mit dem berichteten Einsatz von kognitiven Organisationsstrategien (Organisieren), den metakognitiven Strategien sowie der ressourcenbezogenen Strategie des Rückgriffs auf zusätzliche Literatur. Lehramtsstudierende, die der Überzeugung sind, dass erziehungswissenschaftlicher Wissenserwerb einer vielfältigen, elaborierten Auseinandersetzung mit den Studientexten bedarf (im Gegensatz zu einfachem Durchlesen), berichten demnach häufiger Organisationsstrategien einzusetzen, ihr Lernverhalten häufiger aktiv zu planen und zu regulieren sowie häufiger auf zusätzliche Literatur zurückzugreifen. Zwischen der StEB-Skala *Entwicklung des Wissens* und der LIST-R-Skala *metakognitive Lernstrategien* besteht ein signifikanter positiver

Zusammenhang. Lehramtststudierende, die der Überzeugung sind, dass erziehungswissenschaftliches Wissen sich weiterentwickelt und nicht gleich bleibt, scheinen mehr Wert auf die Planung und Regulation ihres Lernprozesses zu legen, bzw. sie berichten, dieses häufiger zu tun. Zwischen der kognitiven Lernstrategie *Wiederholen* und den beiden Skalen *Absolutheit* und *Einfachheit des Wissens* bestehen signifikante positive Zusammenhänge. Lehramtsstudierende, die der Überzeugung sind, dass erziehungswissenschaftliches Wissen aus Studientexten aus sicheren, diskreten Fakten besteht, die es zu wissen gilt (einfache Wissensstruktur), setzen häufiger Wiederholungsstrategien beim Lernen erziehungswissenschaftlicher Inhalte ein, bzw. berichten, diese häufiger einzusetzen. Auch Lehramtsstudierende, die glauben, erziehungswissenschaftliche Studientexte stellten eine verlässliche Quelle des Wissens dar, deren Inhalte man nicht anzeifeln sollte, berichten Wiederholungsstrategien häufiger einzusetzen. Unsophistizierte Überzeugungen zur *Einfachheit des Wissens* (erziehungswissenschaftliches Wissen als sichere, diskrete zu wissende Fakten) gehen darüber hinaus mit weniger Rückgriff auf zusätzliche Literatur aber mehr Nutzung von Onlineressourcen einher. Die Überzeugung, dass erziehungswissenschaftliches Wissen absolut im Gegensatz zu relativ ist, steht in einem positiven Zusammenhang mit einer expliziten Zeitplanung für das Lernen.

In Abbildung 8.1 auf Seite 292 wird die in der Tabelle 8.7 dargestellte und soeben beschriebene empirisch gefundene Zusammenhangsstruktur noch einmal überblicksartig dargestellt. Auf der linken Seite sind die beiden StEB-Skalen *Einfachheit* und *Absolutheit* abgebildet, bei denen eine starke Zustimmung *unsophistizierten* Überzeugungen zu erziehungswissenschaftlichem Wissen entspricht. Auf der rechten Seite sind dagegen die Skalen *Multimodalität* und *Entwicklung* abgebildet, bei denen eine starke Zustimmung *sophistizierten* Überzeugungen zu erziehungswissenschaftlichem Wissen entspricht. In der Mitte sind die vier Skalen der kognitiven Lernstrategien des LIST-R zu sehen. In Klammern unter der jeweiligen Skalenbezeichnung finden sich die Abkürzungen GW, SW, EW, QW und RW. Diese Abkürzungen zeigen an, aus welchen Items sich die entsprechende StEB-Skala zusammensetzt. Items, die sich beispielsweise auf die Quelle des Wissens (QW) und die Rechtfertigung des Wissens (RW) beziehen, bilden gemeinsam die StEB-Skala Absolutheit des Wissens. Die StEB-Skala Einfachheit des Wissens besteht dagegen aus Items zur Sicherheit (GW) und Struktur des Wissens (SW). Die Multimodalität des Wissenserwerbs setzt sich aus Items zur Rechtfertigung (RW), Sicherheit (GW) und Quelle des Wissens (QW) zusammen, während die Skala Entwicklung des Wissens ausschließlich Items enthält, die sich auf die Entwicklung des Wissens (EW) beziehen.

Die roten Pfeile stehen für hypothesenkonforme, empirisch gefundene negative Zusammenhänge zwischen einer StEB- und einer LIST-R-Skala, die grünen für ebensolche positiven Zusammenhänge. Hypothesenkonform sind solche Zusammenhänge, die bereits auf Basis der Theorie postuliert, in zahlreichen empirischen Studien bestätigt wurden und im nomologischen Netz in Kapitel 2.10 in Abbildung 2.18 (S.130) überblicksartig dargestellt werden. Die blauen Pfeile stellen Zusammenhänge dar, die theoretisch angenommen waren,

aber zu denen in den vorliegenden Daten der zweiten Stichprobe ($N = 891$) keine entsprechenden Zusammenhänge gefunden wurden. Die angenommenen negativen Zusammenhänge zwischen der StEB-Skala *Multimodalität des Wissenserwerbs* und *Wiederholungsstrategien* sowie *Entwicklung des Wissens* und *Wiederholungsstrategien* wurden nicht gefunden. Ebenso wurden auch die erwarteten negativen Zusammenhänge zwischen *Absolutheit des Wissens* mit *metakognitiven Strategien* sowie *Einfachheit des Wissens* mit *metakognitiven Strategien* und *Organisationsstrategien* nicht gefunden.

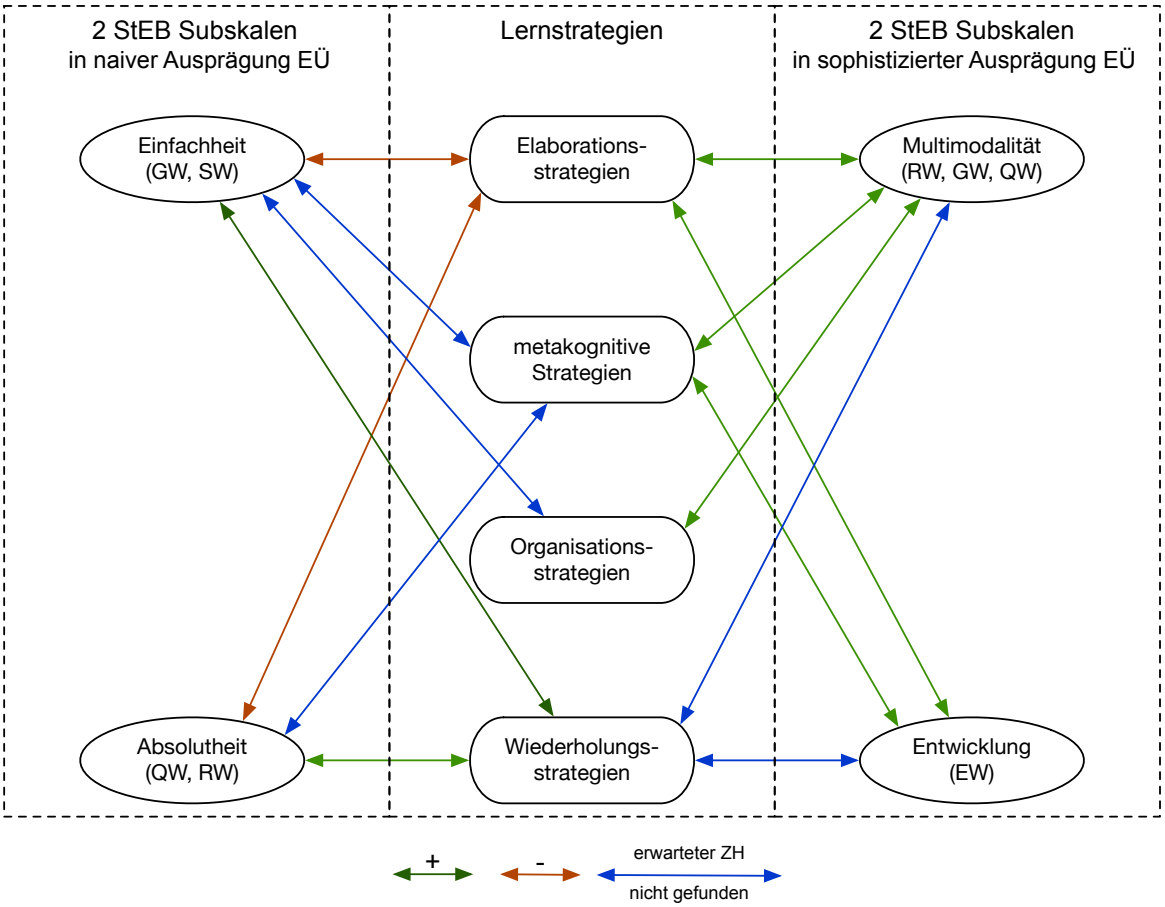


ABBILDUNG 8.1: Empirische Zusammenhangsstruktur zwischen den StEB-Subskalen und Lernstrategien

Bezüglich der Prüfung, ob ein Zusammenhang sich hypothesenkonform verhält oder nicht, musste natürlich die Itemzusammensetzung der StEB-Skalen berücksichtigt werden. Das nomologische Netz, das in Kapitel 2.10 in Abbildung 2.18 (S.130) aufgestellt wurde und mit dem die empirisch gefundenen Zusammenhänge (siehe Abb. 8.1, S. 292) abgeglichen werden müssen, enthält beispielsweise einen positiven Zusammenhang zwischen der sophistizierten Ausprägung Rechtfertigung des Wissens mit den Elaborationsstrategien sowie den metakognitiven Strategien. Demnach verhalten sich die empirischen Zusammenhänge hypothesenkonform, wenn die sophistiziert gepolten StEB-Skalen, die ein Item-Subset zur Rechtfertigung des Wissens enthalten, ebenfalls positive Zusammenhänge mit den LIST-R-Skalen Elaborationsstrategien und metakognitiven Strategien aufweisen, die unsophistiziert

gepolten StEB-Skalen mit Rechtfertigungsitems dagegen negative Zusammenhänge mit diesen aufweisen. Das betrifft sowohl die Skala *Multimodalität* als auch die Skala *Absolutheit*, da beide Skalen Items zur Rechtfertigung des Wissens enthalten. Die Skala Absolutheit des Wissens, bei der hohe Zustimmung unsophistizierte epistemologische Überzeugungen widerspiegelt, weist einen negativen Zusammenhang zu den Elaborationsstrategien auf, der demnach hypothesenkonform ist. Die Skala *Multimodalität* weist einen positiven Zusammenhang zu Elaborationsstragegien und metakognitiven Strategien auf, die in ihrer Richtung demnach ebenfalls hypothesenkonform sind. Auf diese Art wurden alle gefundenen Zusammenhänge auf Hypothesenkonformität mit dem nomologischen Netz abgeglichen und entsprechend in Abbildung 8.1 (S. 292) eingezeichnet.

Aus Abbildung 8.1 (S. 292) lässt sich entnehmen, dass sich viele der angenommenen Zusammenhänge (siehe Kap. 2.10, Abb. 2.18, S.130) zwischen den Skalen des StEB und des LIST-R tatsächlich finden lassen. Die gefundenen Korrelationen entsprechen sowohl in ihrer Richtung als auch in der Stärke des Zusammenhangs den in zahlreichen Studien bereits empirisch gezeigten Zusammenhängen (vgl. Abschnitt 2.8.2, ab S. 96). Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten weisen tatsächlich moderate Zusammenhänge mit der berichteten präferierten Lernstrategieauswahl beim Lernen mit erziehungswissenschaftlichen Studientexten auf. Sophistiziertere Ausprägungen epistemologischer Überzeugungen in allen vier StEB-Dimensionen (*Multimodalität*, *Entwicklung*, *Einfachheit* und *Absolutheit*) gehen mit den tiefenorientierten Lernstrategien *Elaboration*, *metakognitive Lernstrategien* und *Organisation* einher, während unsophistiziertere Ausprägungen epistemologischer Überzeugungen zur *Einfachheit* und *Absolutheit* des erziehungswissenschaftlichen Wissens aus Studientexten mit der oberflächenorientierten Wiederholungsstrategie einhergehen. Grundannahme 4 kann demnach als bestätigt gelten, was wiederum die Gültigkeit der Testwertinterpretationen des StEB zusätzlich stützt.

## 8.4 Unterschiedsprüfungen mittels MANOVA

Zur Prüfung der Grundannahmen 5 bis 8 werden die Unterschiede in den Überzeugungsausprägungen der Lehramstudierenden in Abhängigkeit der Hochschule, der Schulart, des Studienfortschritts und der fachlichen Ausrichtung untersucht.

**Grundannahme 5:** Es bestehen Unterschiede in den Ausprägungen epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten zwischen Lehramtsstudierenden verschiedener Hochschulstandorte.

**Grundannahme 6:** Es bestehen Unterschiede in den Ausprägungen epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten zwischen Lehramtsstudierenden der Studiengänge Real-, Grund-, Sekundar-, Mittel- und Gymnasialschullehramt (Schulart).

**Grundannahme 7:** Es bestehen Unterschiede in den Ausprägungen epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten zwischen Lehramtsstudierenden hoher und niedriger Fachsemester.

**Grundannahme 8:** Es bestehen Unterschiede in den Ausprägungen epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten zwischen Lehramtsstudierenden hoher Fachsemester mit sozialwissenschaftlicher, naturwissenschaftlicher und geisteswissenschaftlicher Ausrichtung. Für Lehramtsstudierende niedriger Fachsemester mit sozialwissenschaftlicher, naturwissenschaftlicher und geisteswissenschaftlicher Ausrichtung finden sich jedoch keine interindividuellen Unterschiede in den Ausprägungen epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen.

Die Grundannahmen 5 bis 8 wurden je mittels multivariater Varianzanalyse (MANOVA) vom Typ 3 geprüft. Typ 3 wurde aufgrund der unbalancierten Zellengrößen verwendet. Als Testwerte wurden zuerst die individuellen Skalenmittelwerte des finalen 26-Item StEB Inventars verwendet, deren Histogramme im Anhang I in den Abbildungen I.1 bis I.4 (ab S. I.1) dargestellt sind. Auch die Histogramme der Faktorwerte sind im Anhang I in den Abbildungen I.5 bis I.8 (ab S. I.5) zu sehen, da die im folgenden beschriebenen Analysen zusätzlich mit Faktorwerten als StEB-Testwerten durchgeführt wurden.

Zur **Prüfung von Grundannahme 5** wurde untersucht, ob Personen sich in Abhängigkeit von der Hochschulzugehörigkeit in ihren Überzeugungsausprägungen zur Absolutheit, Einfachheit, Multimodalität und Entwicklung von erziehungswissenschaftlichem Wissen bezüglich Studientexten unterscheiden.

Vor der Analysen musste jedoch der Datensatz der Validierungsstichprobe (Studie 2) erneut bereinigt werden. Da die MANOVA Mittelwertsunterschiede mehrerer abhängiger Variablen zwischen Gruppen testet, dürfen keine Gruppen vorkommen, die zu gering besetzt sind. Deshalb wurden die Datensätze von den Personen entfernt, die keinen Studienstandort angegeben hatten sowie die Datensätze der nur 14 Lehramtsstudierenden, die an der Universität Potsdam studieren. Außerdem wurden die elf Datensätze von Personen entfernt, die angegeben hatten die Universität Freiburg zu besuchen, obwohl sie in einer Lehrveranstaltung der PH Freiburg befragt worden waren. Zu diesen beiden Standorten lagen zu wenige Daten vor, um sie als eigene Gruppe in die MANOVA mit aufzunehmen.

In Tabelle 8.8 sind die durchschnittlichen Überzeugungsausprägungen der Studierenden nach Hochschule aufgeschlüsselt. Hohe Ausprägungen in den Subskalen *Absolutheit* und *Einfachheit* entsprechen unsophistizierten Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten, während niedrige Ausprägungen sophistizierte Überzeugungen widerspiegeln.

TABELLE 8.8: Mittelwerte und Standardabweichungen der StEB-Überzeugungsausprägungen getrennt nach Hochschule

| Studienstandort/ StEB-Dimension | *M* | *SD* | Studienstandort/ StEB-Dimension | *M* | *SD* |
|---|---|---|---|---|---|
| Uni Passau ($n$ = 402) | | | Uni Regensburg ($n$ = 56) | | |
| Absolutheit | 3.13 | .61 | Absolutheit | 3.03 | .62 |
| Einfachheit | 3.77 | .61 | Einfachheit | 3.62 | .63 |
| Multimodalität | 3.73 | .54 | Multimodalität | 3.72 | .50 |
| Entwicklung | 3.60 | .58 | Entwicklung | 3.53 | .59 |
| Uni Halle ($n$ = 162) | | | Uni Bremen ($n$ = 51) | | |
| Absolutheit | 2.92 | .67 | Absolutheit | 2.94 | .60 |
| Einfachheit | 3.36 | .64 | Einfachheit | 3.05 | .57 |
| Multimodalität | 3.84 | .54 | Multimodalität | 4.01 | .40 |
| Entwicklung | 3.46 | .56 | Entwicklung | 3.45 | .45 |
| PH Freiburg ($n$ = 178) | | | | | |
| Absolutheit | 3.04 | .63 | | | |
| Einfachheit | 3.58 | .63 | | | |
| Multimodalität | 3.85 | .52 | | | |
| Entwicklung | 3.57 | .53 | | | |

Für die Subskalen *Multimodalität* und *Entwicklung* entsprechen hohe Ausprägungen sophistizierten Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten und niedrige unsophistizierten. Die 162 Hallenser Lehramtsstudierenden verfügen mit $M$ = 2.92 ($SD$ = .67) über die durchschnittlich geringste Ausprägung bezüglich der Absolutheit des Wissens, dicht gefolgt von den Bremer Lehramtsstudierenden mit $M$ = 2.94 ($SD$ = .60). Die Gruppe der Bremer Lehramtsstudierenden verfügt über die durchschnittlich geringste Ausprägung in den Überzeugungen zur Einfachheit des Wissens ($M$ = 3.05, $SD$ = .57). Die höchste durchschnittliche Ausprägung zur Absolutheit des Wissens ($M$ = 3.13, $SD$ = .61) und Einfachheit des Wissens ($M$ = 3.77, $SD$ = .61) weisen die Passauer Lehramtsstudierenden auf. Die Gruppe der Bremer Lehramtsstudierenden verfügt mit $M$ = 4.01 ($SD$ = .40) über die durchschnittlich höchste Ausprägung in den Überzeugungen zur Multimodalitiät des Wissenserwerbs. Die Gruppe der Regensburger Lehramtsstudierenden verfügt mit $M$ = 3.72 ($SD$ = .50) über die durchschnittlich niedrigste Ausprägung zur Multimodalität.

Unter Verwendung der Pillai Teststatistik zeigte sich ein signifikanter Einfluss des Hochschulstandortes auf die Überzeugungsausprägungen zur Absolutheit, Einfachheit, Multimodalität und Entwicklung von erziehungswissenschaftlichem Wissen bezüglich Studientexten ($V$ = 0.13, $F$(16,3376) = 6.9, $p$ < .001). Zur Untersuchung der Kontraste wurde je abhängiger Variable, d.h. je Überzeugungsdimension, zunächst eine ANOVA post-hoc berechnet. In Abhängigkeit der Hochschule liegen signifikante Unterschiede in den Testwerten der Befragten bezüglich der Absolutheit des Wissens ($F$(4,844) = 3.73, $p$ < .01, $\eta^2$ = .017), der Einfachheit des Wissens ($F$(4,844) = 23.65, $p$ < .001, $\eta^2$ = .10), der Multimodalität des Wissenserwerbs ($F$(4,844) = 4.34, $p$ < .01, $\eta^2$ = .02) und der Entwicklung des Wissens ($F$(4,844) = 2.43, $p$ < .05, $\eta^2$ = .01) vor.

Der Post-hoc Test (Tukey HSD) zeigte bezüglich der Überzeugung zur Absolutheit des Wissens nur einen einzigen signifikanten Unterschied zwischen den Lehramtsstudierenden der Universität Passau und denen der Universität Halle ($p$ < .01). Wie oben im Zusammenhang mit den Ergebnissen der zugehörigen ANOVA berichtet, werden mit einer Effektstärke von $\eta^2$ = .017 nur 1,7% der Varianz aufgeklärt. Der Effekt ist sehr gering. Weitere Unterschiede in Abhängigkeit des Hochschulstandortes zeigten sich für die Absolutheit des Wissens nicht. Für die Überzeugungen zur Einfachheit des Wissens identifizierte der Post-hoc Test (Tukey HSD) signifikante Unterschiede in der durchschnittlichen Merkmalsausprägung zwischen den Studierenden der Uni Passau und Uni Halle ($p$ < .001), der Uni Passau und PH Freiburg ($p$ < .01), der Uni Passau und Uni Bremen ($p$ < .001) und der Uni Halle und PH Freiburg ($p$ < .05), der Uni Halle und Uni Regensburg ($p$ < .05), der Uni Halle und Uni Bremen ($p$ < .05) sowie der Uni Bremen und der PH Freiburg ($p$ < .001) und der Uni Bremen und der Uni Regensburg ($p$ < .001). Mit einer Effektstärke von $\eta^2$ = .10 werden immerhin 10% der Varianz aufgeklärt. Es handelt sich um einen kleinen Effekt. Für die Überzeugungen zur Multimodalität des Wissens zeigten sich nur zwischen den Studierenden der Uni Passau und Bremen ($p$ > .01) sowie der Uni Bremen und Regensburg ($p$ > .05) signifikante Unterschiede. Der Effekt ist mit $\eta^2$ = .02, d.h. nur 2% aufgeklärter Varianz klein. Trotz der signifikanten ANOVA konnte der paarweise Vergleich keinen signifikanten Unterschied mehr zwischen den Studienstandorten bezüglich der Überzeugungen zur Entwicklung des

Wissens feststellen ($p > .05$). Den folgenden Abbildungen 8.2, 8.4 (S. 299) und 8.6 (S. 300) sind die durchschnittlichen Überzeugungsausprägungen getrennt nach Hochschule zu entnehmen. Die Abbildungen 8.3 (S. 298), 8.5 (S. 299) und 8.7 (S. 301) zeigen die Boxplots der durchschnittlichen Ausprägungen (Skalenmittelwerte) zur Absolutheit, Entwicklung und Multimodalität des erziehungswissenschaftlichen Wissens getrennt nach Hochschule. Die entsprechenden Abbildungen für die Überzeugung zur Entwicklung des Wissens befinden sich im Angang I in den Tabellen I.10 (S. 446) und I.9 (S. 446).

Die Überzeugungen zur Absolutheit des Wissens bewegen sich für alle fünf Gruppen nahe 3 (von 5). Da hohe Ausprägungen hier unsophistizierte Überzeugungen widerspiegeln, verfügen alle Studierenden über relativ unsophistizierte Überzeugungen zur Absolutheit des Wissens. Die Hallenser Studierenden verfügen jedoch im Vergleich zu den Passauer Studierenden über signifikant sophistiziertere Überzeugungen zur Absolutheit des erziehungswissenschaftlichen Wissens (siehe Abb. 8.2 und Abb. 8.3, S. 298). Sie sind durchschnittlich etwas weniger der Überzeugung, dass erziehungswissenschaftliche Studientexte verlässliche Quellen für erziehungswissenschaftliches Wissen darstellen, deren Inhalte man ohne Hinterfragen glauben sollte.

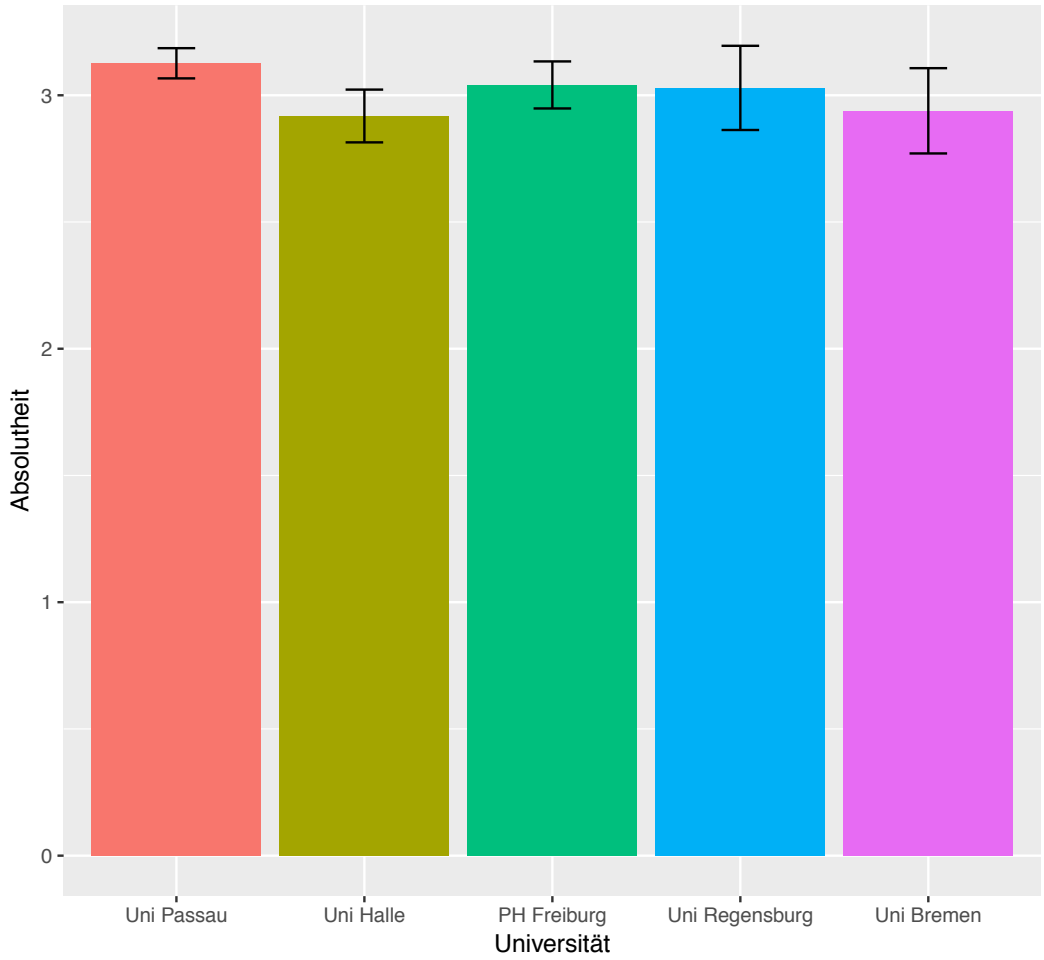


ABBILDUNG 8.2: Durchschnittliche Ausprägung in der Überzeugung zur Absolutheit des Wissens nach Hochschule

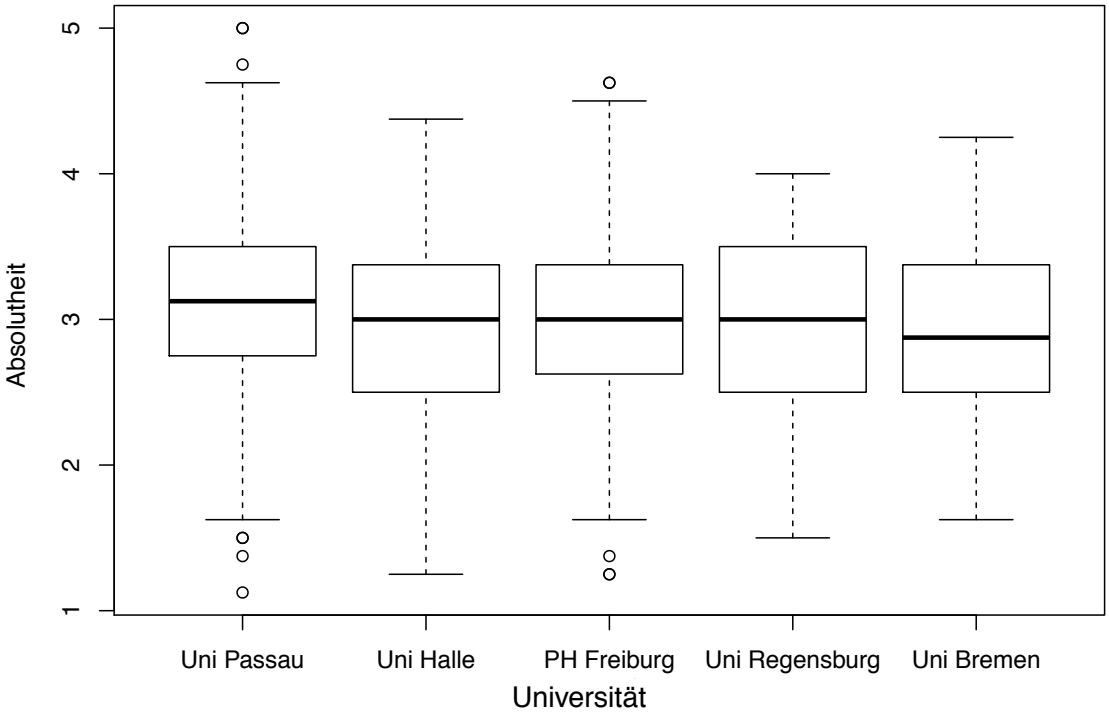


ABBILDUNG 8.3: Boxplot der individuellen Skalenmittelwerte zur Absolutheit des Wissens nach Hochschule

Die durchschnittlichen Überzeugungen zur Einfachheit des Wissens sind für alle fünf Gruppen über 3 (von 5) und folglich ebenfalls hoch, was einer relativ unsophistizierten Ausprägung entspricht. Die Passauer Lehramtsstudierenden verfügen über die durchschnittlich unsophistiziertesten Überzeugungen zur Einfachheit des erziehungswissenschaftlichen Wissens. Sie sind signifikant stärker der Überzeugung, dass erziehungswissenschaftliches Wissens aus sicheren, diskreten zu wissenden Fakten besteht, als die Hallenser, Freiburger und Bremer Lehramtsstudierenden (siehe Abb 8.4, S. 299). Die Bremer Studierenden verfügen über die durchschnittlich sophistiziertesten Überzeugungen zur Einfachheit des erziehungswissenschaftlichen Wissens. Der Überzeugung, dass erziehungswissenschaftliches Wissens aus sicheren, diskreten zu wissenden Fakten besteht, ist bei den Bremer Studierenden signifikant geringer ausgepägt als bei den Lehramtsstudierenden der vier anderen Hochschulen (siehe Abb 8.4, S. 299). Die Hallenser Lehramtstudierenden verfügen durchschnittlich über sophistiziertere Überzeugungen zur Einfachheit des erziehungswissenschaftlichen Wissens als die Lehramtstudierenden der Uni Passau, PH Freiburg und Uni Regensburg.

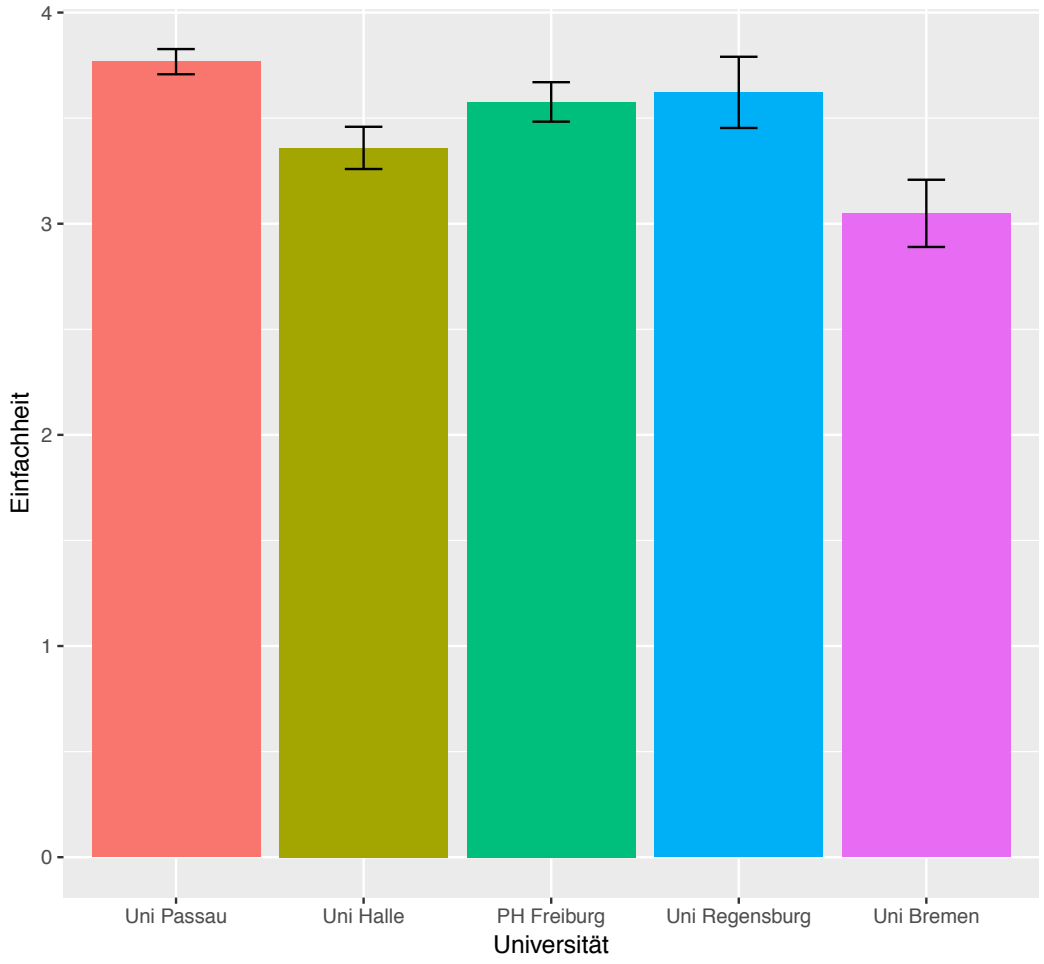


ABBILDUNG 8.4: Durchschnittliche Ausprägung in der Überzeugung zur Einfachheit des Wissens nach Hochschule

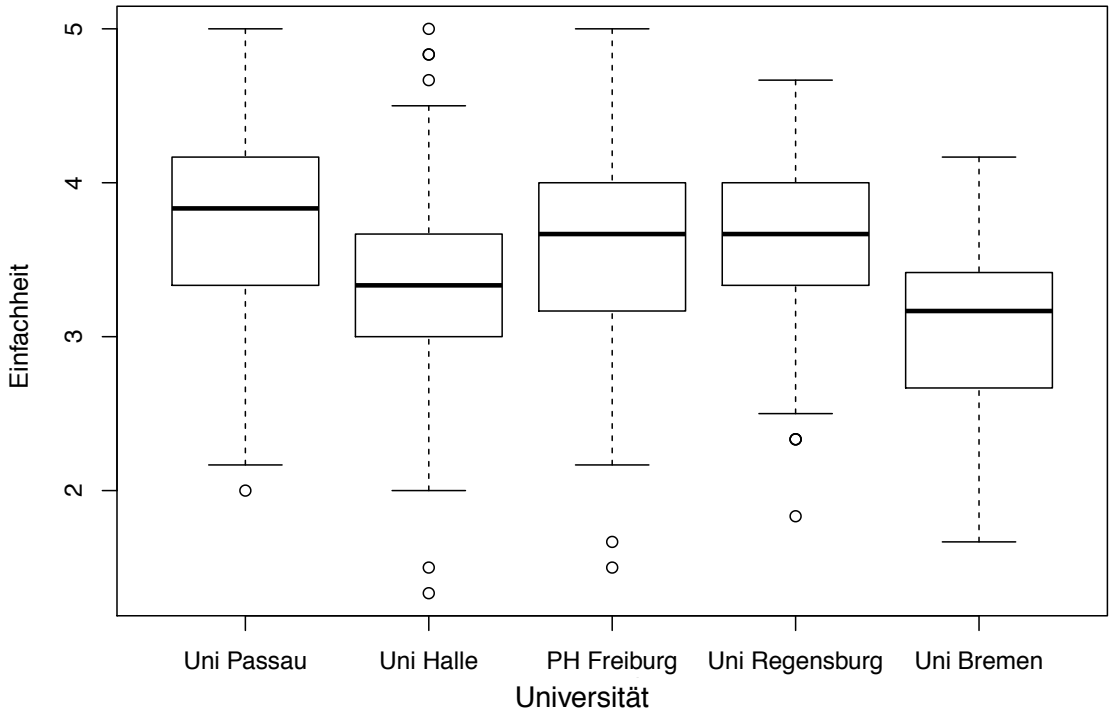


ABBILDUNG 8.5: Boxplot der individuellen Skalenmittelwerte zur Einfachheit des Wissens nach Hochschule

Die durchschnittlichen Überzeugungen zur Multimodalität des erziehungswissenschaftlichen Wissenserwerbs sind für alle fünf Gruppen mit Werten über 3 (von 5) relativ hoch ausgeprägt. Das entspricht einer relativ sophistizierten Überzeugung. Die Bremer Lehramtsstudierenden verfügen im Vergleich zu den Passauer und Regensburger Studierenden über signifikant sophistiziertere Überzeugungen zur Multimodalität des Wissenserwerbs (siehe Abb. 8.6 und Abb. 8.7, S. 301). Sie sind also stärker der Überzeugung, dass der erziehungswissenschaftliche Wissenserwerb einer vielfältigen, elaborierten Auseinandersetzung mit erziehungswissenschaftlichen Studientexten bedarf.

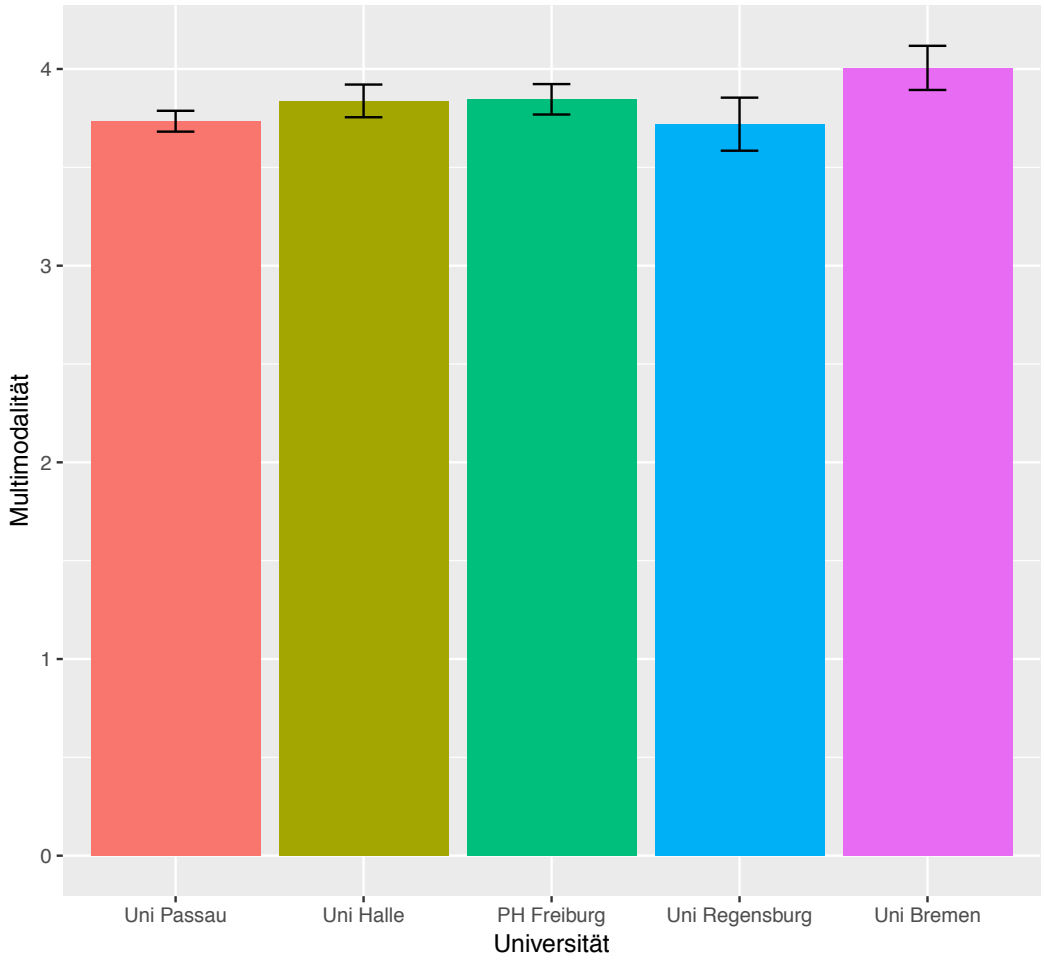


ABBILDUNG 8.6: Durchschnittliche Ausprägung in der Überzeugung zur Multimodalität des Wissenserwerbs nach Hochschule

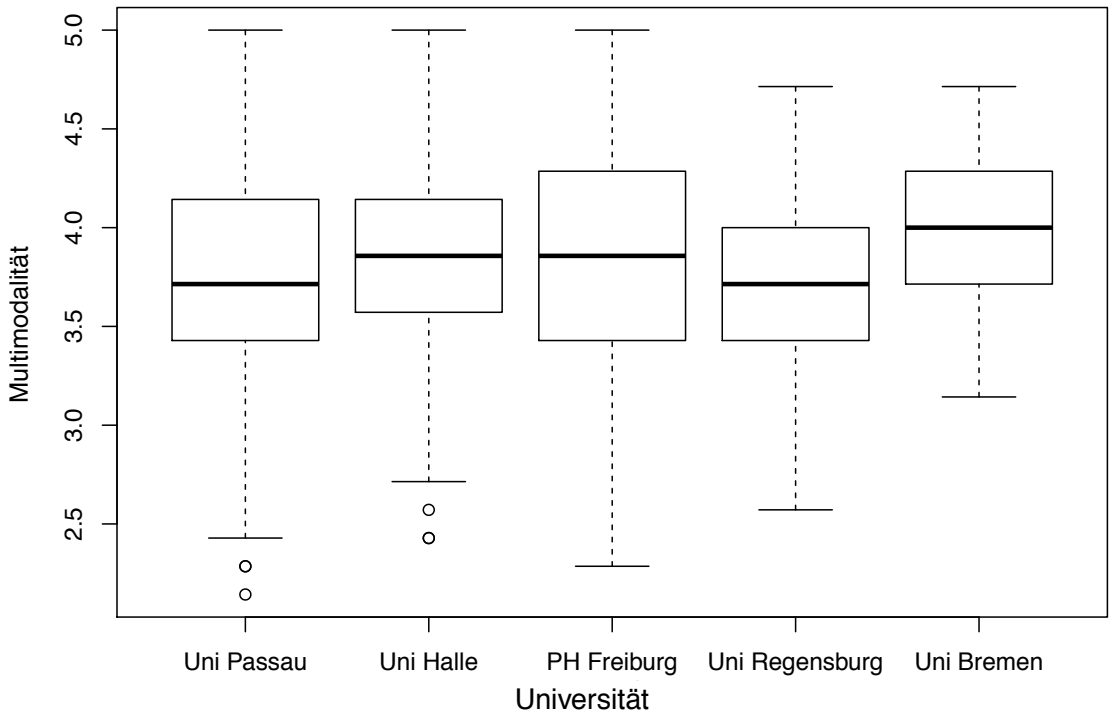


ABBILDUNG 8.7: Boxplot der individuellen Skalenmittelwerte zur Multimodalität des Wissenserwerbs nach Hochschule

Die Überzeugungen zur Entwicklung des Wissens sind für alle fünf Hochschulen durchschnittlichen sehr ähnlich ausgeprägt und mit Werten über 3 (von 5) relativ hoch. Das entspricht einer relativ sophistizierten Überzeugung zur Entwicklung des erziehungswissenschaftlichen Wissens (siehe Anhang I Abb. I.9 und Abb. I.10, S. 446).

Zusammenfassend lässt sich festhalten, dass interindividuelle Unterschiede in den Überzeugungsausprägungen zu erziehungswissenschaftlichem Wissen in Abhängigkeit von der jeweiligen Hochschulzugehörigkeit der Lehramtsstudierenden gefunden werden konnten. Einschränkend ist jedoch anzumerken, dass die gefundenen Effekte gering sind. Die Grundannahme 5 kann als bestätigt gelten, was wiederum die Gültigkeit der Testwertinterpretationen stützt.

Zur **Prüfung von Grundannahme 6** wurde untersucht, ob Personen sich in Abhängigkeit der studierten Schulart in ihren Überzeugungsausprägungen zur Absolutheit, Einfachheit, Multimodalität und Entwicklung von erziehungswissenschaftlichem Wissen bezüglich Studientexten unterscheiden.

Die Daten der Befragten, die Gesamt- und Hauptschullehramt als Schulart angegeben hatten, wurden von den Analysen ausgeschlossen, da sie insgesamt nicht häufig genug vertreten waren, um in der MANOVA eine eigene Gruppe darzustellen. Auch Datensätze von Personen, die keine Schulart angegeben hatten, wurden entfernt. Die Studierenden, die als Schulart Förderschule und die Studierenden, die als Schulart Grund- und Förderschule angegeben hatten, wurden zu einer Gruppe zusammengefasst.

In Tabelle 8.9 sind die durchschnittlichen Überzeugungsausprägungen der Studierenden nach Schulart dargestellt. Hohe Ausprägungen in den Subskalen *Absolutheit* und *Einfachheit* entsprechen unsophistizierten Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten, während niedrige Ausprägungen sophistizierte Überzeugungen widerspiegeln. Für die Subskalen *Multimodalität* und *Entwicklung* entsprechen hohe Ausprägungen sophistizierten Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten und niedrige unsophistizierten wissensbezogenen Überzeugungen.

TABELLE 8.9: Mittelwerte und Standardabweichungen der StEB Überzeugungsausprägungen getrennt nach Schulart

| Schulart/ StEB-Dimension | *M* | *SD* | Schulart/ StEB-Dimension | *M* | *SD* |
|---|---|---|---|---|---|
| Grundschule ($n$ = 332) | | | Förderschule ($n$ = 32) | | |
| Absolutheit | 3.17 | .63 | Absolutheit | 2.79 | .54 |
| Einfachheit | 3.68 | .64 | Einfachheit | 3.20 | .63 |
| Multimodalität | 3.77 | .51 | Multimodalität | 4.01 | .43 |
| Entwicklung | 3.55 | .55 | Entwicklung | 3.62 | .39 |
| Realschule ($n$ = 98) | | | Gymnasium ($n$ = 169) | | |
| Absolutheit | 3.11 | .54 | Absolutheit | 2.79 | .64 |
| Einfachheit | 3.66 | .54 | Einfachheit | 3.34 | .68 |
| Multimodalität | 3.71 | .53 | Multimodalität | 3.83 | .55 |
| Entwicklung | 3.48 | .55 | Entwicklung | 3.58 | .61 |
| Mittelschule ($n$ = 100) | | | Sekundarschule ($n$ = 113) | | |
| Absolutheit | 3.09 | .62 | Absolutheit | 3.07 | .57 |
| Einfachheit | 3.83 | .62 | Einfachheit | 3.57 | .62 |
| Multimodalität | 3.74 | .57 | Multimodalität | 3.87 | .52 |
| Entwicklung | 3.59 | .57 | Entwicklung | 3.58 | .53 |

Die Studierenden, die Gymnasial- und Förderschullehramt studieren, verfügen mit $M$ = 2.79 über die durchschnittlich geringste Überzeugung zur Absolutheit des Wissens, während diese bei den Studierenden des Grundschullehramts mit $M$ = 3.17 durchschnittlich am höchsten ausgeprägt ist. Die Studierenden des Förderschullehramts verfügen über die durchschnittlich geringste Überzeugung zur Einfachheit des Wissens und wiederum höchste zur Multimodalität des Wissenserwerbs sowie zur Entwicklung des Wissens (siehe Abb. 8.9, S. 302). Die Überzeugung zur Einfachheit des Wissens ist bei den Lehramtsstudierenden der Mittelschule mit $M$ = 3.83 ($SD$ = .62) durchschnittlich am höchsten ausgeprägt. Über die durchschnittlich geringsten Überzeugungen zur Multimodalität des Wissenserwerbs und Entwicklung des Wissens verfügen die Studierenden des Realschullehramts (siehe Abb. 8.9, S. 302).

Die MANOVA (Pillai Teststatistik) zeigte einen signifikanten Unterschied in den Überzeugungsausprägungen zur Absolutheit, Einfachheit, Multimodalität und Entwicklung von erziehungswissenschaftlichem Wissen bezüglich Studientexten in Abhängigkeit von der Schulart ($V$ = 0.10, $F$(20,3544) = 4.8, $p$ < .001). Bezüglich den Ausprägungen zur Absolutheit des Wissens ($F$(5,899) = 11.00, $p$ < .001, $\eta^2$ = .06), Einfachheit des Wissens ($F$(5,899) = 13.42, $p$ < .001, $\eta^2$ = .07) und Multimodalität des Wissenserwerbs ($F$(5,899) = 2.54, $p$ < .05, $\eta^2$ = .01) bestehen signifikante Unterschiede zwischen den Lehramtsstudierenden der verschiedenen Schularten (ANOVA je Überzeugungsdimension).

Im Post-hoc Test (Tukey HSD) zeigten sich für die Überzeugung zur Absolutheit des Wissens signifikante Unterschiede zwischen der Gruppe der Studierenden des Gymnasiallehramts und der Gruppe des Grund- ($p$ < .001), Real- ($p$ < .001), Mittel- ($p$ < .001) und Sekundarschullehramts ($p$ < .01). Die Studierenden des Förderschullehramt und des Grundschullehramts unterscheiden sich ebenfalls signifikant ($p$ < .05) in ihren Überzeugungen zur Absolutheit des Wissens. Mit einer Effektstärke von $\eta^2$ = .06 werden 5,7% der Varianz aufgeklärt. Auch für die Überzeugung zur Einfachheit des Wissens identifizierte der Post-hoc Test signifikante Unterschiede zwischen der Gruppe des Gymnasiallehramts und den Gruppen des Grundschullehramts ($p$ < .001), Realschullehramts ($p$ < .001) und Mittelschullehramts ($p$ < .001). Auch die Lehramtsstudierenden für Förderschule unterscheiden sich in ihrer Überzeugung zur Einfachheit des Wissens signifikant von den Gruppen für Grundschule ($p$ < .01), Realschule ($p$ < .01) und Mittelschule ($p$ < .001). Zwischen den Lehramtsstudierenden für Sekundar- und Mittelschule liegt ebenfalls ein signifikanter Unterschied in der Ausprägung zur Einfachheit des Wissens vor ($p$ < .01). Auch hier liegt nur ein kleiner Effekt von $\eta^2$ = .07 vor (7,0% aufgeklärte Varianz). Trotz des signifikanten F-Wertes zeigten sich für die Überzeugung zur Multimodalität des Wissensserwerbs im Post-hoc Test keine Unterschiede in Abhängigkeit der Schulart.

Die Abbildungen 8.8 und 8.10 (S. 304 und 306) zeigen die Ausprägungen der Überzeugung zur Absolutheit und Einfachheit des Wissens für die Gruppen der Lehramtsstudierenden der unterschiedlichen Schularten. In den Abbildungen 8.9 und 8.11 (S. 305 und 306) sind die Boxplots der Skalenmittelwerte für die Überzeugungen zur Absolutheit und Einfachheit des Wissens getrennt nach Schulart dargestellt. Im Anhang I ab Seite 447 sind die entsprechenden Abbildungen für die übrigen Überzeugungsdimensionen zu sehen.

Die Überzeugungen zur Absolutheit des Wissens bewegen sich für alle sechs Gruppen nahe 3 (von 5). Da hohe Ausprägungen hier unsophistizierte Überzeugungen widerspiegeln, verfügen alle Studierenden über relativ unsophistizierte Überzeugungen zur Absolutheit des Wissens (siehe Abb. 8.8, S. 304). Die Lehramtsstudierenden für Gymnasium verfügen jedoch im Vergleich zu den Grund- , Real-, Mittel- und Sekundarschulstudierenden über signifikant sophistiziertere Überzeugungen zur Absolutheit des erziehungswissenschaftlichen Wissens (siehe Abb. 8.8, S. 304 und Abb. 8.9, S. 305). Sie sind durchschnittlich weniger der Überzeugung, dass erziehungswissenschaftliche Studientexte verlässliche Quellen für erziehungswissenschaftliches Wissen darstellen, deren Inhalte man ohne Hinterfragen glauben sollte. Die Förderschulstudierenden verfügen über eine durchschnittlich ähnlich sophistizierte Überzeugung zur Absolutheit des Wissens wie die Studierenden für Gymnasiallehramt. Signifikant wird der Unterschied jedoch nur zwischen den Gruppen Förderschule und Grundschule. Die Grundschullehramtsstudierenden verfügen im Vergleich zu den Studierenden für Förderschule signifikant stärker über die unsophistizierte Überzeugung, dass erziehungswissenschaftliche Studientexte verlässliche Quellen für erziehungswissenschaftliches Wissen darstellen, deren Inhalte man ohne Hinterfragen glauben sollte (siehe Abb. 8.8, S. 304 und Abb. 8.9, S. 305).

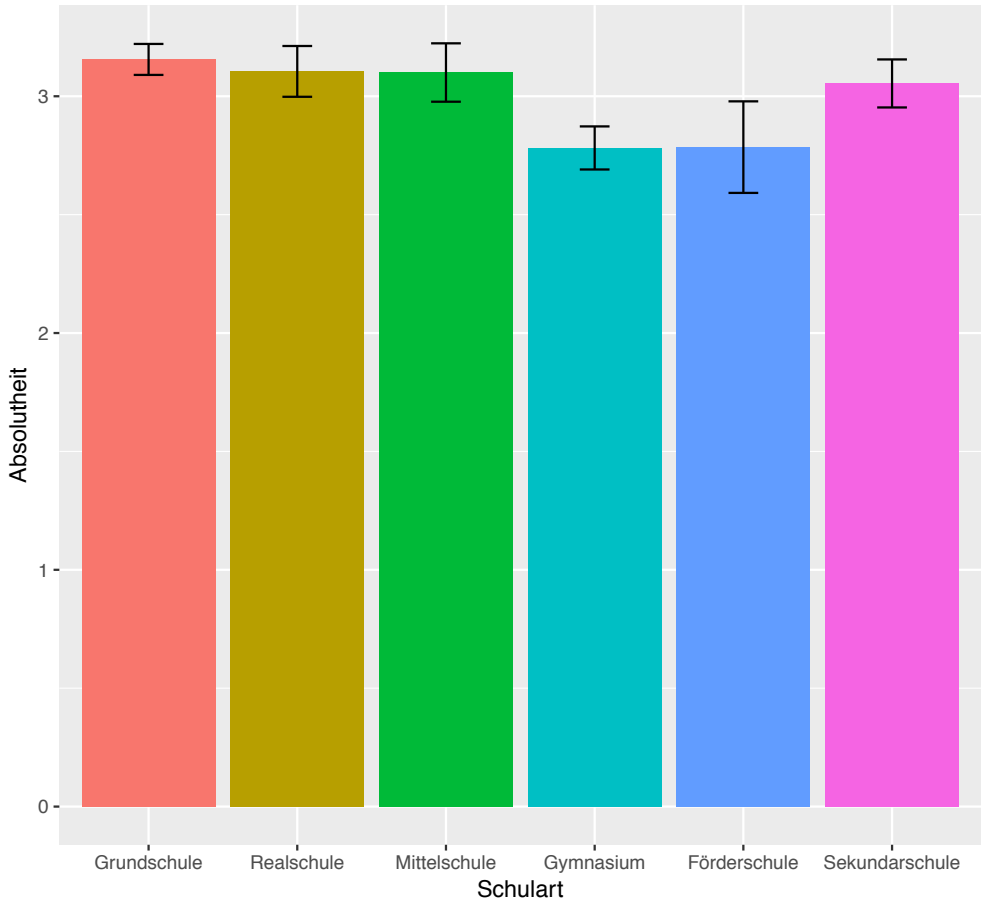


ABBILDUNG 8.8: Durchschnittliche Ausprägung in der Überzeugung zur Absolutheit des Wissens nach Schulart

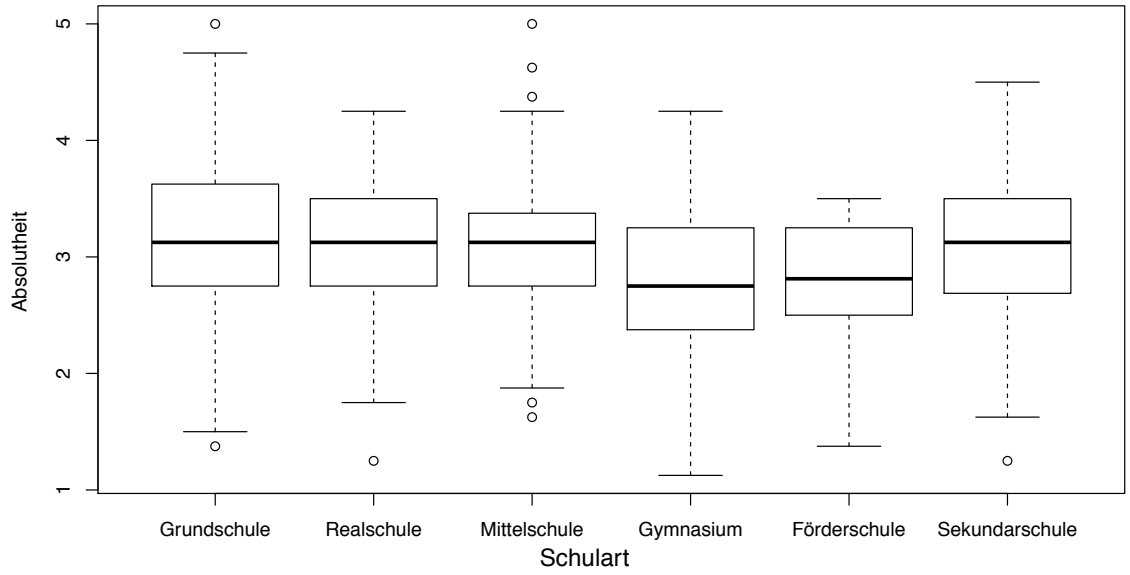


ABBILDUNG 8.9: Boxplot der individuellen Skalenmittelwerte zur Absolutheit des Wissens nach Schulart

Die Lehramtsstudierenden für Gymnasium und Förderschule verfügen im Vergleich zu den anderen Gruppen über die sophistiziertesten Überzeugungen zur Einfachheit des Wissens (siehe Abb. 8.10 und Abb. 8.11, S. 306). Die durchschnittlichen Überzeugungen zur Einfachheit des Wissens sind allerdings für alle sechs Gruppen mit durchschnittlichen Werten über 3 (von 5) als relativ unsophistiziert zu beschreiben. Die Studierenden für Gymnasium und Förderschule sind im Vergleich zu den Gruppen für Grund-, Real- und Mittelschule signifikant weniger davon überzeugt, dass erziehungswissenschaftliches Wissen aus sicheren, diskreten zu wissenden Fakten besteht. Die Sekundarschulstudierenden verfügen über signifikant sophistiziertere Überzeugungen zur Einfachheit des Wissens als die Mittelschulstudierenden, d.h. sind weniger der Ansicht, dass erziehungswissenschaftliches Wissen aus sicheren, diskreten zu wissenden Fakten besteht (siehe Abb. 8.10 und Abb. 8.11, S. 306).

Bezüglich Grundannahme 6 lässt sich festhalten, dass interindividuelle Unterschiede in den Überzeugungsausprägungen zu erziehungswissenschaftlichem Wissen in Abhängigkeit von der jeweilig studierten Schulart der Lehramtsstudierenden gezeigt werden konnten, wenn auch mit geringen Effekten. Die Grundannahme 6 kann somit als bestätigt gelten, was wiederum die Gültigkeit der Testwertinterpretationen stützt.

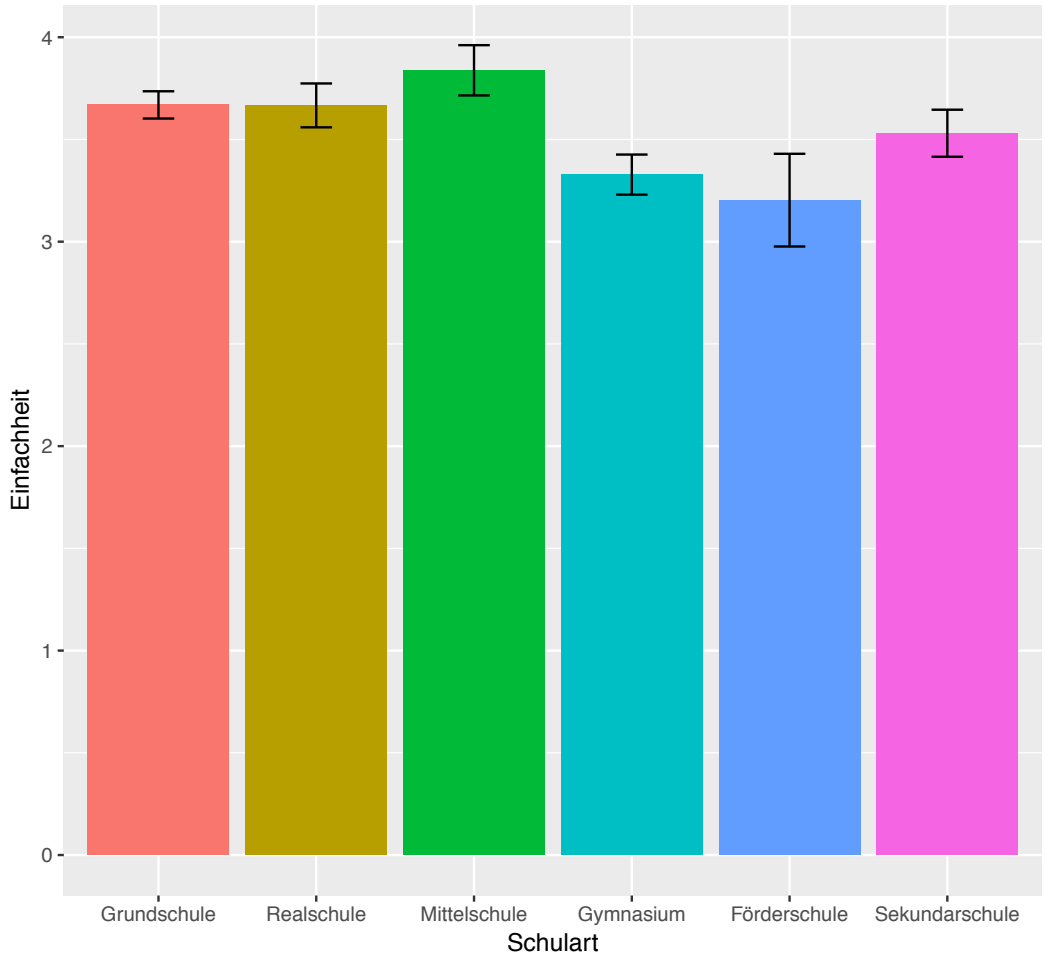


ABBILDUNG 8.10: Durchschnittliche Ausprägung in der Überzeugung zur Einfachheit des Wissens nach Schulart

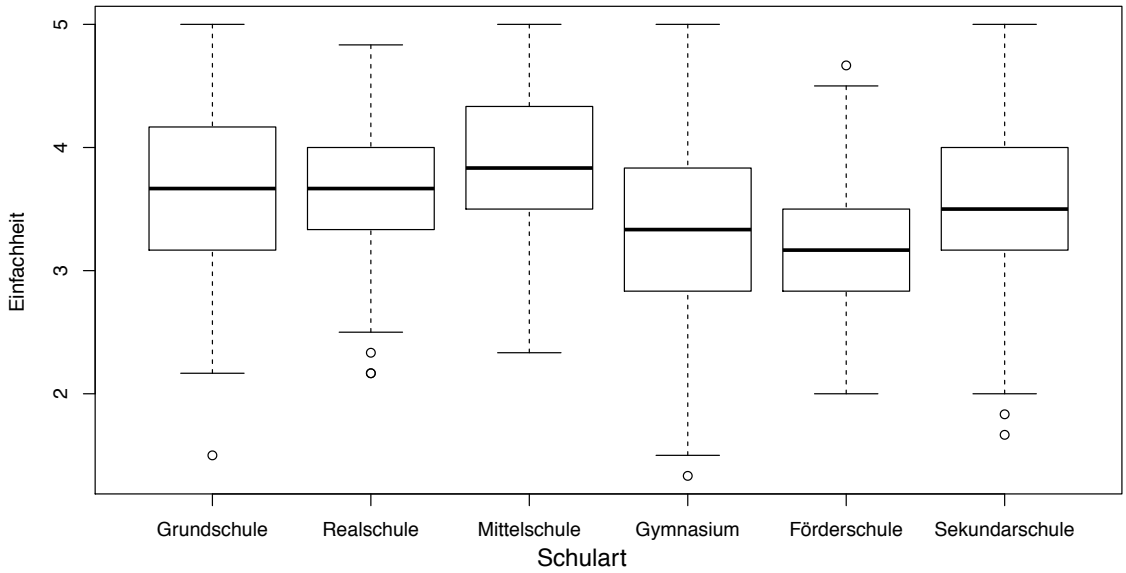


ABBILDUNG 8.11: Boxplot der individuellen Skalenmittelwerte zur Einfachheit des Wissens nach Schulart

Zur **Prüfung von Grundannahme 7** wurde getestet, ob Studienanfänger, d.h. Personen im ersten oder zweiten Fachsemester, sich in ihren Überzeugungsausprägungen zur Absolutheit, Einfachheit, Multimodalität und Entwicklung von erziehungswissenschaftlichem Wissen bezüglich Studientexten von Personen im fortgeschrittenen Fachstudium unterscheiden.

In Tabelle 8.10 sind die durchschnittlichen Überzeugungsausprägungen der Studienanfänger und fortgeschrittenen Studierenden zu sehen. Hohe Ausprägungen in den Subskalen *Absolutheit* und *Einfachheit* entsprechen unsophistizierten Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten, während niedrige Ausprägungen sophistizierte Überzeugungen widerspiegeln. Für die Subskalen *Multimodalität* und *Entwicklung* entsprechen hohe Ausprägungen dagegen sophistizierten Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten und niedrige unsophistizierten. Die Studienanfänger im ersten und zweiten Fachsemester verfügen im Durchschnitt über eine höhere Überzeugung zur Einfachheit des Wissens ($M$ = 3.69, $SD$ = .64) und Entwicklung des Wissens ($M$ = 3.60, $SD$ = .55) als die fortgeschrittenen Studierenden höherer Semester. Die Überzeugungen zur Multimodalität des Wissenserwerbs sind zwischen beiden Gruppen fast gleich hoch und die Überzeugungen zur Absolutheit des Wissens unterscheiden sich nicht (siehe Tab. 8.10, S. 307).

TABELLE 8.10: Mittelwerte und Standardabweichungen der StEB Überzeugungsausprägungen getrennt nach Studienfortschritt

| Studienfortschritt/ StEB-Dimension | *M* | *SD* |
|---|---|---|
| Studienanfänger (n = 465) | | |
| Absolutheit | 3.05 | .61 |
| Einfachheit | 3.69 | .64 |
| Multimodalität | 3.80 | .53 |
| Entwicklung | 3.60 | .55 |
| fort. Fachstudierende (n = 384) | | |
| Absolutheit | 3.05 | .65 |
| Einfachheit | 3.49 | .65 |
| Multimodalität | 3.79 | .54 |
| Entwicklung | 3.50 | .57 |

Die MANOVA (Pillai Teststatistik) identifizierte einen signifikanten Unterschied in den Überzeugungsausprägungen zur Absolutheit, Einfachheit, Multimodalität und Entwicklung von erziehungswissenschaftlichem Wissen bezüglich Studientexten in Abhängigkeit vom Studienfortschritt ($V$ = 0.03, $F(4,844)$ = 7.7, $p < .001$). Allerdings besteht nur bezüglich der

Überzeugung zur Einfachheit des Wissens ($F(1,847) = 20.70$, $p < .001$, $\eta^2 = .02$) und der Überzeugungen zur Entwicklung des Wissens ($F(1,847) = 6.37$, $p < .05$, $\eta^2 = .007$) je ein signifikanter Unterschied zwischen den Studienanfängern und den fortgeschrittenen Fachsemestern.

Die Überzeugung zur Einfachheit des Wissens ist bei beiden Gruppen mit durchschnittlichen Ausprägungen zwischen 3 und 4 (von 5) stark vertreten, was unsophistizierten Überzeugungen entspricht. Die Studienanfänger verfügen über die signifikant unsophistiziertere Überzeugung zur Einfachheit des Wissens. Sie sind stärker als die fortgeschrittenen Studierenden davon überzeugt, dass Wissen aus sicheren, diskreten zu wissenden Fakten besteht. Mit 2,4% aufgeklärter Varianz ist der Effekt gering. Gleichzeitig verfügen die Studienanfänger im Vergleich zu den fortgeschrittenen Studierenden über die signifikant sophistiziertere Überzeugung zur Entwicklung des Wissens, da sie stärker der Überzeugung sind, dass sich das in Studentexten dargestellte erziehungswissenschaftliche Wissen über die Zeit verändert. Der Effekt ist mit nur 0,7% aufgeklärter Varianz allerdings sehr gering.

Zusammenfassend lässt sich festhalten, dass bezüglich der Überzeugungsdimension *Einfachheit des Wissens* ein Unterschied zwischen Studienanfängern und fortgeschrittenen Studierenden gefunden werden konnte. Der Effekt dieses interindividuellen Unterschieds ist sehr gering. Ein weiterer gefundener Unterschied in der Dimension *Entwicklung des Wissens* ist zu gering und nicht bedeutsam. Die Ergebnisse zur Grundannahme 7 zeichnen insgesamt kein eindeutiges Bild. Grundannahme 7 kann deshalb nicht angenommen werden.

Zur **Prüfung von Grundannahme 8** sollte geprüft werden, ob bei fortgeschrittenen Studierenden (Fachsemester > 3) in Abhängigkeit von der fachlichen Ausrichtung (sozialwissenschaftlich, geisteswissenschaftlich, naturwissenschaftlich), Unterschiede in den Überzeugungsausprägungen zur Absolutheit, Einfachheit, Multimodalität und Entwicklung von erziehungswissenschaftlichem Wissen bestehen. Es wurde angenommen, dass solche Unterschiede bei Studienanfängern nicht auftreten sollten. Für die Einteilung in die verschiedenen Gruppen nach fachlicher Ausrichtung wurde zunächst anhand der angegebenen Studienfächer (spätere Unterrichtsfächer) eine Kodierung vorgenommen. Es wurde einerseits zwischen Lehramtsstudierenden unterschieden, deren Ausrichtung über alle Fächer hinweg konsequent naturwissenschaftlich (NW), geisteswissenschaftlich (GW) oder sozialwissenschaftlich (SW) ist. Da andererseits häufig gemischte Fachprofile vorliegen, wurde hier unterschieden zwischen dominant naturwissenschaftlicher (dNW), geisteswissenschaftlicher (dGW) oder sozialwissenschaftlicher (dSW) Ausrichtung. Auch die ausgeglichen gemischtfachlichen Profile, bei denen keine der fachlichen Ausrichtungen vorherrschte, wurden kodiert (natur/geist; natur/sozial; geist/sozial; natur/geist/sozial). Es war jedoch nicht möglich, den Studienfortschritt (Studienanfänger vs. fortgeschrittenes Fachstudium) bei der Prüfung von Grundannahme 8 mitzuberücksichtigen, da die Zellenbelegung zu klein wird, wenn diese Variable als Kovariate mit in die MANOVA aufgenommen wird.

Tabelle 8.11 zeigt die Einteilung der Fächer in die Bereiche der Geistes-, Natur- und Sozialwissenschaft. Hier ist auch zu entnehmen, wie häufig die einzelnen Fächer unter den $N = 849$ Lehramtsstudierenden vorkommen. Die häufigste Geisteswissenschaft ist das Fach Deutsch, das von 486 Personen studiert wird. 440 Lehramtsstudierende haben Mathematik als Fach angegeben. Mathematik ist damit das häufigste naturwissenschaftliche Fach. Sozialwissenschaftliche Fächer sind deutlich weniger vertreten als geistes- und naturwissenschaftliche.

TABELLE 8.11: Zuordnung der Fächer zu Bereichen und Häufigkeiten der Fächer ($N = 849$)

| Geistes-wissenschaft | Probanden | Natur-wissenschaft | Probanden | Sozial-wissenschaft | Probanden |
|---|---|---|---|---|---|
| Deutsch | 486 | Mathematik | 440 | Sport | 182 |
| Englisch | 230 | Geographie | 122 | Sozialkunde | 61 |
| Kunst | 160 | Biologie | 59 | Arbeitslehre | 44 |
| Geschichte | 158 | Sachunterricht | 43 | SPF | 36 |
| Religion | 106 | Physik | 20 | AKG | 34 |
| Musik | 91 | Chemie | 19 | EZW | 15 |
| Ethik | 28 | Technik | 14 | Politik | 11 |
| Französich | 28 | AWT | 13 | Psychologie | 2 |
| Latein | 12 | Astronomie | 8 | Soziologie | 1 |
| Philosophie | 6 | Informatik | 7 | | |
| Spanisch | 4 | Erdkunde | 6 | | |
| Italienisch | 3 | | | | |
| Russisch | 1 | | | | |

AWT Arbeit Wirtschaft Technik, AKG Alltagskultur und Gesundheit, EZW Erziehungswissenschaft, SPF sonderpädagogische Fächer

Im Folgenden wurde demnach nur geprüft, ob Unterschiede in den epistemologischen Überzeugungen in Abhängigkeit von der fachlichen Ausrichtung auftreten. Der Studienfortschritt wurde hier nicht mitberücksichtigt.

In Tabelle 8.12 auf Seite 311 sind die durchschnittlichen Überzeugungsausprägungen[4] der Studierendengruppen mit unterschiedlicher fachlicher Ausrichtung zu sehen. Die Lehramtsstudierenden mit konsequent naturwissenschaftlichem Fächerprofil (NW) verfügen mit $M$ = 2.93 ($SD$ = .57) über die durchschnittlich geringste Ausprägung in der Absolutheit des Wissens. Über eine genauso geringe, durchschnittliche Ausprägung ($M$ = 2.93 ($SD$ = .67) verfügt die Gruppe GW der konsequent geisteswissenschaftlich ausgerichteten Lehramtsstudierenden. Die höchste Ausprägung zur Absolutheit des Wissens liegt bei Lehramtsstudierenden mit konsequent sozialwissenschaftlichem Fächerprofil (SW) vor ($M$ = 3.20 ($SD$ = .44). Dieser Gruppe gehören allerdings nur 16 Personen an. Die höchste Ausprägung in den Überzeugungen zur Einfachheit des Wissens erreicht mit $M$ = 3.89 ($SD$ = .52) die Gruppe mit dominant sozialwissenschaftlichem Fächerprofil (dSW), über die geringste verfügt mit $M$ = 3.35 ($SD$ = .73) dagegen die Gruppe mit ausgeglichen natur-, geistes- und sozialwissenschaftlichem Fächerprofil (NWGWSW).

Die höchste durchschnittliche Ausprägung zur Multimodalität des Wissenserwerbs erreichen mit je $M$ = 3.86 die Gruppen mit konsequent geisteswissenschaftlichem (GW) sowie die mit gemischt natur- und sozialwissenschaftlichem Fächerprofil (NWSW). Die dominant sozialwissenschaftliche Gruppe (dSW) verfügt dagegen über die durchschnittlich geringste Ausprägung zur Multimodalität des Wissenserwerbs. Die höchste Ausprägung zur Entwicklung des Wissens hat mit $M$ = 3.62 ($SD$ = .56) die dominant geisteswissenschaftliche Gruppe (dGW), über die geringste verfügt dagegen die dominant naturwissenschaftliche Gruppe (dNW, textitM = 3.44, $SD$ = .56).

Die MANOVA (Pillai Teststatistik) zeigte einen signifikanten Unterschied in den epistemologischen Überzeugungen zu erziehungswissenschaftlichem Wissen bezüglich Studientexten zwischen den Gruppen fachlicher Ausrichtung ($V$ = 0.08, $F(36, 3356)$ = 1.96, $p < .001$). Die Untersuchung der Kontraste mittels ANOVA je Überzeugungsdimension resultierte nur für die Einfachheit des Wissens in einem signifikanten Unterschied in Abhängigkeit der fachlichen Ausrichtung ($F(9{,}839)$ = 3.75, $p < .001$, $\eta^2$ = .04). Der Effekt ist mit 3,8% aufgeklärter Varianz klein. Es bestehen signifikante Unterschiede in den Überzeugungen zur Einfachheit des Wissens zwischen der dominant geisteswissenschaftlichen Gruppe und der konsequent geisteswissenschaftlichen Gruppe ($p < .01$). In den Überzeugungen zur Einfachheit des Wissens unterscheidet sich außerdem die dominant sozialwissenschaftliche Gruppe signifikant von der konsequent geisteswissenschaftlichen Gruppe ($p < .05$) sowie die gemischtfachliche Gruppe von der dominant sozialwissenschaftlichen Gruppe ($p < .05$).

[4] Hohe Ausprägungen in den Subskalen *Absolutheit* und *Einfachheit* entsprechen unsophistizierten Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten, während niedrige Ausprägungen sophistizierte Überzeugungen widerspiegeln. Für die Subskalen *Multimodalität* und *Entwicklung* entsprechen hohe Ausprägungen dagegen sophistizierten Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten und niedrige unsophistizierten.

TABELLE 8.12: Mittelwerte und Standardabweichungen der StEB-Überzeugungsausprägungen getrennt nach fachlicher Ausrichtung

| Fächerprofil/ StEB-Dimension | *M* | *SD* | Fächerprofil/ StEB-Dimension | *M* | *SD* |
|---|---|---|---|---|---|
| Naturwiss. Profil (NW) ($n$ = 67) | | | Dominant sozialwiss. (dSW) ($n$ = 29) | | |
| Absolutheit | 2.93 | .57 | Absolutheit | 3.13 | .53 |
| Einfachheit | 3.56 | .57 | Einfachheit | 3.89 | .52 |
| Multimodalität | 3.71 | .59 | Multimodalität | 3.67 | .52 |
| Entwicklung | 3.46 | .62 | Entwicklung | 3.57 | .56 |
| Geisteswiss. Profil (GW) ($n$ = 144) | | | Gemischt natur/geist (NWGW) ($n$ = 124) | | |
| Absolutheit | 2.93 | .67 | Absolutheit | 3.06 | .69 |
| Einfachheit | 3.44 | .62 | Einfachheit | 3.55 | .71 |
| Multimodalität | 3.86 | .52 | Multimodalität | 3.70 | .51 |
| Entwicklung | 3.57 | .57 | Entwicklung | 3.56 | .62 |
| Sozialwiss. Profil (SW) ($n$ = 16) | | | Gemischt natur/sozial (NWSW) ($n$ = 36) | | |
| Absolutheit | 3.20 | .44 | Absolutheit | 3.09 | .48 |
| Einfachheit | 3.57 | .43 | Einfachheit | 3.77 | .56 |
| Multimodalität | 3.71 | .34 | Multimodalität | 3.86 | .55 |
| Entwicklung | 3.49 | .34 | Entwicklung | 3.47 | .36 |
| Dominant naturwiss. (dNW) ($n$ = 68) | | | Gemischt geist/natur (GWNW) ($n$ = 51) | | |
| Absolutheit | 3.07 | .50 | Absolutheit | 3.01 | .63 |
| Einfachheit | 3.47 | .67 | Einfachheit | 3.59 | .71 |
| Multimodalität | 3.71 | .51 | Multimodalität | 3.81 | .50 |
| Entwicklung | 3.44 | .56 | Entwicklung | 3.58 | .52 |
| Dominant geisteswiss. (dGW) ($n$ = 285) | | | Gemischt natur/geist/sozial (NWGWSW) ($n$ = 29) | | |
| Absolutheit | 3.13 | .65 | Absolutheit | 3.00 | .58 |
| Einfachheit | 3.71 | .64 | Einfachheit | 3.35 | .73 |
| Multimodalität | 3.84 | .54 | Multimodalität | 3.83 | .58 |
| Entwicklung | 3.62 | .56 | Entwicklung | 3.45 | .45 |

Abbildung 8.12 zeigt die Ausprägungen der Überzeugung zur Einfachheit des Wissens für die verschiedenen Gruppen fachlicher Ausrichtung. In Abbildung 8.13 (S. 313) sind die Boxplots der Skalenmittelwerte für die Überzeugungen zur Einfachheit des Wissens getrennt nach fachlicher Ausrichtung dargestellt. Im Anhang I ab Seite 449 sind die entsprechenden Abbildungen für die übrigen Überzeugungsdimensionen zu sehen. Die Lehramtsstudierenden mit dominant sozialwissenschaftlicher Ausrichtung (dSW) sind im Vergleich zu den gemischtfachlichen (NWGWSW) und den konsequent geisteswissenschaftlichen Lehramtsstudierenden (GW) signifikant stärker der unsophistizierteren Ansicht, dass Wissen in erziehungswissenschaftlichen Studientexten aus sicheren, diskreten zu wissenden Fakten besteht. Die Gruppe dSW verfügt über die durchschnittlich unsophistiziertesten Überzeugungen zur Einfachheit des erziehungswissenschaftlichen Wissens. Die Gruppe NWGWSW verfügt über die sophistiziertesten Überzeugungen zur Einfachheit des Wissens, d.h. die gemischtfachlichen Lehramtsstudierenden sind am stärksten davon überzeugt, dass erziehungswissenschaftliches Wissen aus unsicheren, untereinander verbundenen Konzepten besteht (siehe Abb. 8.12 und Abb. 8.13, S. 313). Die dominant geisteswissenschaftlichen Lehramtsstudierenden (dGW) verfügen im Vergleich zu den konsequent geisteswissenschaftlichen (SW) über signifikant unsophistiziertere, höhere Überzeugungen zur Einfachheit des Wissens (siehe Abb. 8.12 und Abb. 8.13, S. 313).

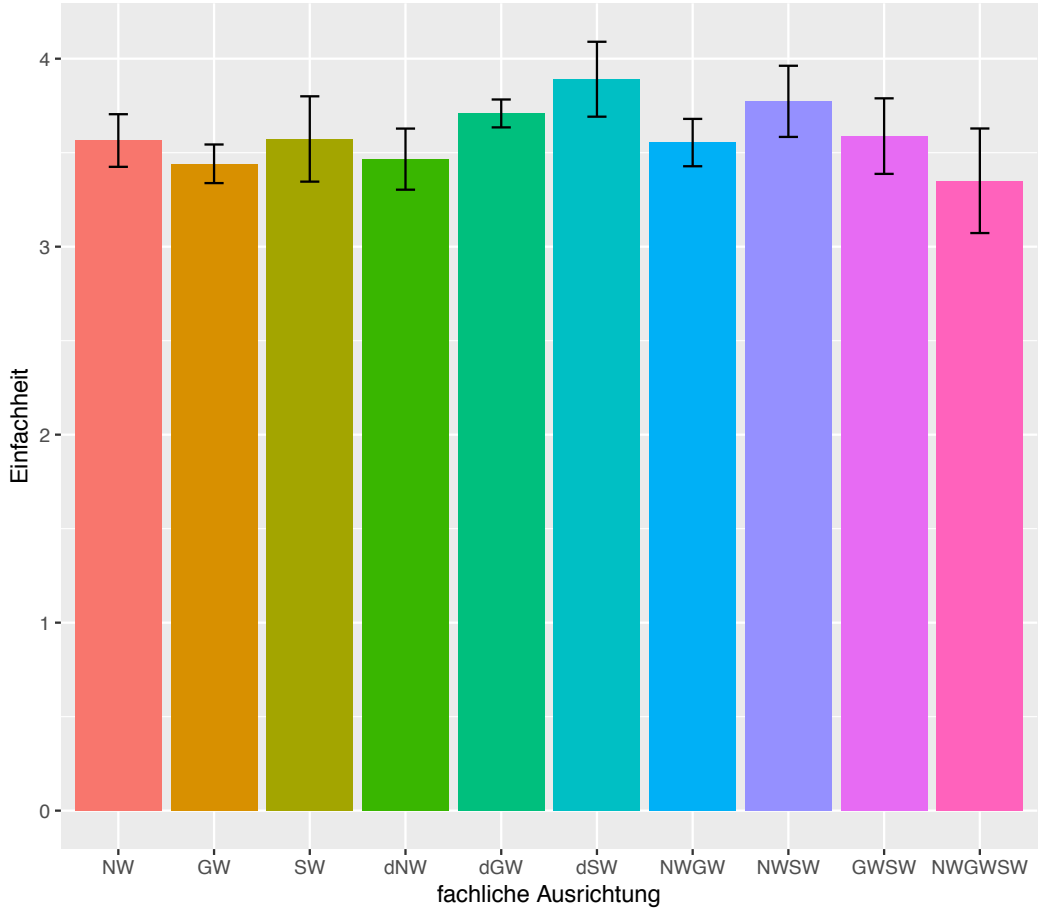


ABBILDUNG 8.12: Durchschnittliche Ausprägung in der Überzeugung zur Einfachheit des Wissens nach fachlicher Ausrichtung

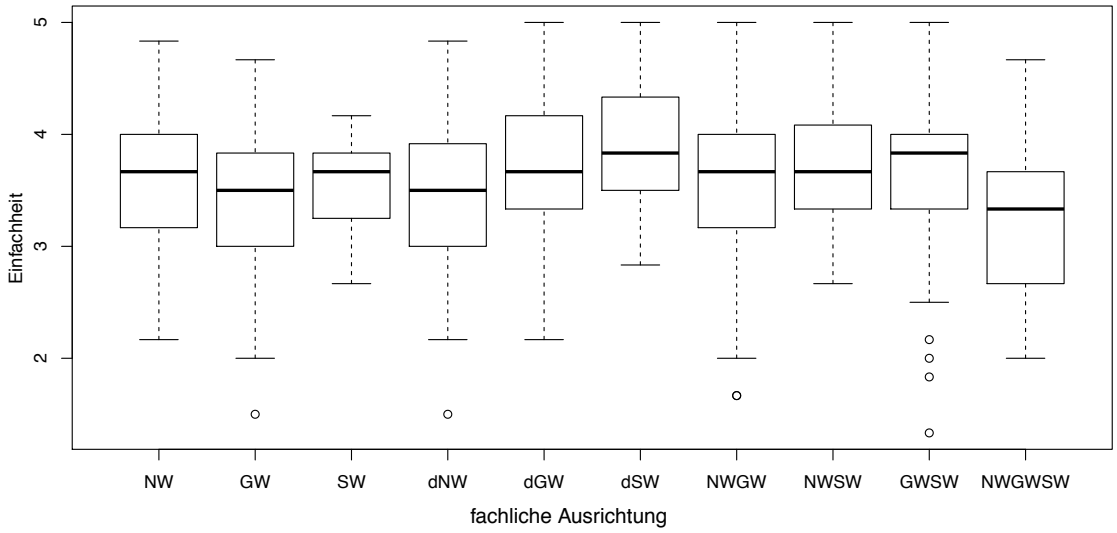


ABBILDUNG 8.13: Boxplot der individuellen Skalenmittelwerte zur Einfachheit des Wissens nach fachlicher Ausrichtung

Über die Grundannahme 8 lassen sich auf Basis dieser Analysen keine Aussagen treffen, da der Studienfortschritt in der Prüfung der Überzeugungsunterschiede zwischen den Gruppen mit unterschiedlicher fachlicher Ausrichtung nicht berücksichtigt werden konnte. Grundannahme 8 kann deshalb weder angenommen noch verworfen werden.

Die Prüfung der Grundannahmen 5 bis 8 wurde statt mit Skalenmittelwerten zusätzlich mit Faktorwerten als Testwerten vorgenommen. Dies führte für alle vier Analysen insofern zu den gleichen Ergebnissen, als dass ohne Ausnahme die gleichen signifikanten Unterschiede identifiziert wurden. Die Effektstärken fielen bei den Analysen anhand der Faktorwerte fast durchgängig etwas höher aus.

Die abgeschlossene Prüfung der acht Grundannahmen stellt verschiedene empirische Evidenzen zur Verfügung, auf deren Basis die Gültigkeit der Testwertinterpretationen des StEB Inventars im Folgenden zusammenfassend beurteilt werden soll.

## 8.5 Zusammenfassende Beurteilung der Evidenzen und Testwertinterpretationen

Zur Prüfung der Grundannahmen 1 und 2 wurde als Evidenzquelle die interne Struktur des Tests sowie der Testinhalt herangezogen. Es wurde auf Ergebnisse der CFA zurückgegriffen, die Informationen über die Faktorstruktur und die Parameterschätzungen lieferte. Die Evidenzen zeigen deutlich, dass die Grundannahme 1 angenommen werden kann, da die Antworten im StEB auf die verschiedenen Überzeugungsausprägungen zurückzuführen sind. Einschränkend wirken hier zwei Items, die für die Erfassung der jeweiligen Überzeugungen weniger geeignet sind und deshalb weiterer Prüfung bedürfen.

Im Rahmen der Prüfung von Grundannahme 2 konnte aufgezeigt werden, dass die aus Studie 1 resultierenden Überzeugungsdimensionen einerseits stabil sind, da sie sich in Studie 2 replizieren ließen. Andererseits wurde auch gezeigt, dass die vier entstandenen Überzeugungsdimensionen sich im Hinblick auf das Zielkonstrukt inhaltlich sinnvoll beschreiben lassen. Darüber hinaus fügen sich die vier entstandenen Dimensionen sinnvoll in die von Hofer und Pintrich (1997) theoretisch angenommenen, übergeordneten Bereiche der *nature of knowing* und *nature of knowledge* ein. Auch Grundannahme 2 kann somit als bestätigt gelten, da das System resultierender Überzeugungsdimensionen zum erziehungswissenschaftlichen Wissen aus Studientexten mit der theoretischen Konzeptionalisierung nach Hofer und Pintrich (1997) vereinbar ist.

Grundannahme 3 postuliert moderate Zusammenhänge zwischen den Überzeugungsdimensionen des StEB. Zur Prüfung dieser Annahme wurden die Kovarianzen zwischen den vier Faktoren analysiert, die von der CFA zu Verfügung gestellt werden. Wie sich zeigte, verhalten sich alle bestehenden Zusammenhänge konstruktkonform zueinander, weshalb auch Grundannahme 3 angenommen werden kann.

Zur Prüfung der Grundannahme 4 wurden die korrelativen Beziehung zwischen den vier Skalen des StEB Inventars mit dem berichteten Lernstrategieeinsatz (LIST-R, Wild, 2011) beim Lernen mit erziehungswissenschaftlichen Studientexten analysiert. Die empirisch vorliegenden Beziehungen zwischen den StEB-Skalen und den LIST-R-Skalen wurden mit der im Rahmen des nomologischen Netzes aufgestellten Zusammenhangsstruktur abgeglichen. Wie sich zeigte, gehen sophistiziertere epistemologische Überzeugungen zum erziehungswissenschaftlichen Wissen hypothesenkonform mit tiefenorientierten Lernstrategien und unsophistiziertere Überzeugungen mit Oberflächenstrategien einher. Viele der angenommenen Zusammenhänge wurden in Stärke und Richtung auch zwischen den StEB- und LIST-R-Skalen empirisch gefunden, während gleichzeitig keine Zusammenhänge vorliegen, die mit dem Theoriegefüge unvereinbar wären. Grundannahme 4 gilt demnach ebenfalls als bestätigt.

Die Grundannahmen 5 bis 8 postulierten Unterschiede in den Überzeugungsausprägungen der vier Dimensionen in Abhängigkeit der Hochschule, der Schulart, der Fachsemester und der fachlichen Ausrichtung in Kombination mit den Fachsemestern. Sie wurden mittels MANOVAs geprüft. Bezüglich Grundannahme 5 und 6 wurden wie angenommen

bei den Lehramtsstudierenden Unterschiede in den Überzeugungsausprägungen zu erziehungswissenschaftlichem Wissen in Abhängigkeit von der jeweiligen Hochschulzugehörigkeit und Schulart gefunden. Die Grundannahmen 5 und 6 können angenommen werden, da die postulierten Unterschiede gefunden wurden. Einschränkend ist jedoch festzuhalten, dass die gefundenen Unterschiede nur geringe Effektstärken aufweisen. Grundannahme 7 kann nicht angenommen werden. Die Studienanfänger verfügen zwar im Vergleich zu den fortgeschrittenen Studierenden hypothesenkonform über signifikant unsophistiziertere Überzeugungen zur Einfachheit des Wissens, allerdings ist der Effekt äußerst gering. Der gefundene Unterschied in der Entwicklung des Wissens, ist so gering, dass er nicht bedeutsam ist. In den anderen beiden Überzeugungsdimensionen wurde kein Unterschied festgestellt. Über Grundannahme 8 lassen sich auf Basis der vorliegenden Daten keine Aussagen treffen, da der Studienfortschritt bei der Prüfung der Überzeugungsunterschiede zwischen den Gruppen mit unterschiedlicher fachlicher Ausrichtung nicht berücksichtigt werden konnte. Die Zellengrößen wurden bei Berücksichtigung des Studienfortschritts (Studienanfänger vs. fortgeschrittene Fachstudierende) als Kovariate zu gering. Deshalb wurde nur geprüft, ob Unterschiede in den epistemologischen Überzeugungen zwischen den Gruppen mit unterschiedlicher fachlicher Ausrichtung vorkommen. Insgesamt sind die gefundenen interindivuellen Unterschiede in den Überzeugungsausprägungen nicht sehr groß und verfügen nur über sehr geringe Effektstärken. Eine Fehlerquelle bei der Prüfung der Mittelwertsunterschiede ist die Abhängigkeit des Signifikanztests von der Stichprobengröße. Bei großen Gruppen und gleichzeitig sehr geringen Unterschieden, wie z.B. dem in der Überzeugung zur Entwicklung des Wissens zwischen den Studienanfängern und fortgeschrittenen Studierenden (Grundannahme 7), ist davon auszugehen, dass dieser Unterschied nur aufgrund der großen Gruppen signifikant wurden.

Die Grundannahmen 1 bis 4 ließen sich bestätigen und stützen demnach die Gültigkeit der Testwertinterpretationen. Die Grundannahmen 6 und 7 werden eingeschränkt angenommen und stützen die Testwertinterpretationen demnach auch, wenn auch schwächer. Die Grundannahme 7 stützt die Gültigkeit der Testwertinterpretationen nicht, während Grundannahmen 8 nicht geprüft werden konnte. Zusammengefasst deuten die gesammelten Evidenzen, die sich auf den Testinhalt, die interne Teststruktur und die Beziehungen des Zielkonstrukts zu anderen Variablen beziehen, auf eine vorläufige Angemessenheit und Gültigkeit der beiden intendierten Testwertinterpretationen. Die Tatsache, dass die MANOVAs mittels Faktorwerten als StEB-Testwerten zu den gleichen Ergebnissen führten, wie mit Skalenmittelwerten als Testwerten, stellt zusätzlich ein starkes Argument dafür dar, dass die Verwendung von Skalenmittelwerten als Testwerte legitim ist. Das StEB Inventar scheint sich folglich dazu zu eignen, die bei Lehramtsstudierenden aktuell vorherrschenden Ausprägungen in den Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten mehrdimensional zu ermitteln und erlaubt zudem einen intraindividuellen und interindividuellen Vergleich. Das StEB Inventar kann folglich dazu genutzt werden, Aussagen darüber zu treffen, ob ein Lehramtsstudierender oder eine Lehramtsstudierende im Vergleich zu einem oder einer anderen über weniger sophistizierte oder aber sophistiziertere

epistemologische Überzeugung zu erziehungswissenschaftlichem Wissen aus Studientexten (in den verschiedenen Dimensionen) verfügt. Über einen wiederholten Einsatz des StEB bei denselben Lehramtsstudierenden ist es außerdem möglich, individuelle Veränderungen in den Ausprägungen epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten (zu den verschiedenen Dimensionen) zu identifizieren.

# 9 Diskussion

Das zentrale Ziel dieser Arbeit besteht in der Entwicklung eines Tests, dem StEB Inventar, das die mehrdimensionale Erfassung der disziplinspezifischen epistemologischen Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen aus Studientexten erlaubt. Die damit verbundene Fragestellung dieser Arbeit bezieht sich darauf, ob sich die epistemologischen Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen aus Studientexten in der Faktorstruktur des StEB konform zum Theoriegefüge abbilden lassen. Diese Instrumententwicklung, die auf Basis der klassischen Testtheorie erfolgte, wurde in den vorangegangenen Kapiteln vorgestellt. In diesem Kapitel wird zunächst der Ablauf dieser Instrumententwicklung zusammengefasst. In Unterkapitel 9.2 wird der Bezug zwischen dem StEB Inventar und dem entwicklungspsychologischen Ansatz epistemologischer Überzeugungen aufgezeigt, bevor in Unterkapitel 9.3 die normative Betrachtung epistemologischer Überzeugungen diskutiert wird. Unterkapitel 9.4 beleuchtet die Limitationen des Entwicklungsprozesses dieser Arbeit und des StEB Inventars. Die Arbeit schließt mit einer Darstellung der forschungspragmatischen und praktischen Implikationen des aus dieser Arbeit resultierenden StEB Inventars.

## 9.1 Zusammenfassung

Im Kapitel zum theoretischen Hintergrund dieser Arbeit wurde in den vielseitigen Forschungsbereich epistemologischer Überzeugungen zunächst eingeführt, indem ein Überblick über die Forschungsentwicklung und die zentralen Forschungsdiskurse gegeben wurde. Neben der Konzeptualisierung Hofer und Pintrichs (1997), die den theoretischen Kern dieser Arbeit darstellt, wurden vier weitere bedeutende Konzeptionen epistemologischer Überzeugungen vorgestellt, die einerseits für den gesamten Forschungsbereich, aber im speziellen auch für diese Instrumententwicklung von Bedeutung sind. Das umfasst das *Scheme of Intellectual and Ethical Development* (Perry, 1970), das *Reflective Judgment Model* von King und Kitchener (1994), Kuhns *Development of Epistemological Understanding* (1991) und das *Epistemological Belief System* von Schommer (1990). Dem folgte eine Darstellung des Entwicklungsverlaufs epistemologischer Überzeugungen, welche die Natur der Entwicklung sowie die Auslöser von Entwicklung und Veränderung beleuchtet. In diesem Zusammenhang wurde das *Integrative Model for Personal Epistemology* (Bendixen & Rule, 2004) vorgestellt, das die Mechanismen der Veränderung epistemologischer Überzeugungen beschreibt. Als zentrale Forschungsdiskurse wurde im Anschluss auf die Mehrdimensionalität sowie die Domänenspezifität und Kontextabhängigkeit eingegangen, die im *Theory of Integrated*

*Domains in Epistemology Framework* (Muis et al., 2006) vereint werden. Dem folgte die Verortung epistemologischer Überzeugungen in Theorien der Kognition und Metakognition. Da sich die Relevanz epistemologischer Überzeugungen von Lehramtsstudierenden aus ihrer Beziehung zum Lehr- und Lernprozess ergibt, wurde auf diese Beziehung umfassend eingegangen. Anhand der Modelle von Hofer (2001) und Schommer-Aikins (2004) sowie zahlreicher Forschungsbefunde wurde aufgezeigt, welche Verbindungen die epistemologischen Überzeugungen zu den verschiedenen lernrelevanten Faktoren aufweisen und wie sie den Lernprozess und das Lehren beeinflussen. Einen Schwerpunkt bildete dabei die Verbindung der wissensbezogenen Überzeugungen zu den Lernstrategien, weshalb hierzu auch die bekannte Lernstrategiekonzeption von Pintrich et al. (1991) skizziert wurde. Dem folgte eine Darstellung der zentralen Erhebungsinstrumente epistemologischer Überzeugungen, wobei der Fokus auf der quantitativen Erhebung und den Limitationen der bestehenden Skalen lag.

Der theoretische Hintergrund wurde in einer Synthese zusammengeführt, die die Positionierung des Konstrukts epistemologischer Überzeugungen in dieser Arbeit aufzeigte und überblicksartig zusammenfasste, auf welche theoretischen Modelle und Konzeptionen welche Aspekte der Definition des Konstrukts in dieser Arbeit zurückgehen. Die zuvor dargestellten Beziehungen zwischen den epistemologischen Überzeugungen und den Lernstrategien wurden in einem nomologischen Netz festgehalten, das die Zusammenhangsstruktur zwischen beiden Konstrukten abbildet. In der Synthese wurde auch die besondere Bedeutung epistemologischer Überzeugungen für den Lehr-Lernprozess von angehenden Lehrkräften aufgezeigt sowie die Relevanz der Instrumententwicklung dieser Arbeit argumentativ hergeleitet.

Aus der theoretischen Grundlage dieser Arbeit geht hervor, dass es sich bei den epistemologischen Überzeugungen um ein mehrdimensionales Konstrukt handelt. Dieses System von Überzeugungen ist nach Hofer und Pintrich (1997) theorieähnlich organisiert und betrifft übergeordnet die individuellen Überzeugungen zum Wissen und Wissensprozess (vgl. Hofer, 2004a, 2005, 2016; Hofer & Pintrich, 1997). Die Kerndimensionen sind Überzeugungen zur *Quelle des Wissens*, *Rechtfertigung des Wissensprozesses*, *Stabilität* bzw. *Sicherheit des Wissens* und *Struktur* bzw. *Einfachheit des Wissens* (Hofer & Pintrich, 1997). Als weitere Dimension wurde die *Entwicklung des Wissens* identifiziert Conley et al. (2004). Die Überzeugungen zur Stabilität, Struktur und Entwicklung des Wissens können dem übergeordneten Bereich *Beschaffenheit des Wissens* zugeordnet werden, während die Überzeugungen zur Quelle und Rechtfertigung des Wissens zum Bereich *Beschaffenheit des Wissensprozesses* gehören. Individuen verfügen sowohl über erstens allgemeine, übergreifende nicht-akademische Überzeugungen (Alltag, Arbeitsumfeld etc.), zweitens allgemeine akademische Überzeugungen als auch über drittens disziplinspezifische akademische Überzeugungen (Muis et al., 2006). Die wissensbezogenen Überzeugungen einer Person werden im Kontext – im instruktionalen, akademischen und soziokulturellen – aktiviert und von diesem beeinflusst (Hofer, 2005, 2006b; Muis et al., 2006). Sie entstehen über die gesamte Lebenspanne (Muis et al., 2006) in dynamischer Interaktion des Individuums mit dem alltäglichen sozialen, kulturellen und akademischen Umfeld (Baxter Magolda, 2004; Belenky et al., 1986; Bendixen & Rule, 2004;

Hofer & Pintrich, 1997) und werden von Erziehung, Bildung, Kultur sowie akademischen Erfahrungen beeinflusst (King & Kitchener, 1994; Kuhn, 1991; Kuhn & Pearsall, 2000; Muis et al., 2006), was zu einer Weiterentwicklung der allgemeinen nicht-akademischen, der allgemeinen akademischen sowie der disziplinspezifischen epistemologische Überzeugungen einer Person führt (vgl. IM, Bendixen & Rule, 2004, TIDE, Muis et al., 2006 sowie Hofer, 2004a). Mit Eintritt in das Bildungssystem und mit fortschreitenden Bildungserfahrungen entwickelt ein Individuum in Auseinandersetzung mit verschiedenen Unterrichtsfächern (in der Schule) und akademischen Disziplinen (in der Universität) parallel jeweils verschiedene disziplinspezifische epistemologische Überzeugungen (vgl. TIDE, Muis et al., 2006). Mit zunehmender akademischer Erfahrung und Spezialisierung in einer bestimmten Disziplin nimmt der Einfluss dieser disziplinspezifischen wissensbezogenen Überzeugungen auf die allgemeinen akademischen Überzeugungen zu (vgl. TIDE, Muis et al., 2006). Die Entwicklung der drei Überzeugungssysteme (allgemein nicht-akademisch, allgemein akademisch, disziplinspezifisch akademisch), die jeweils Überzeugungen zu allen Dimensionen enthalten, verläuft von weniger sophistizierten bis hin zu sophistizierten individuellen wissensbezogenen Überzeugungen (vgl. IM, Bendixen & Rule, 2004; Rule & Bendixen, 2010 sowie auch Hofer, 2005; King & Kitchener, 1994; Kuhn, 1991; Kuhn & Pearsall, 2000; Perry, 1970), wobei rückläufige Entwicklungen auch auftreten und verursacht werden können (Bendixen & Rule, 2004; Muis et al., 2006; Rule & Bendixen, 2010). In der unsophistizierteren Perspektive wird Wissen als feststehend, unveränderlich, aus einfachen, unverbundenen Fakten bestehend, von Autoritäten vermittelt und als zweifelsfrei wahr oder falsch klassifizierbar betrachtet, ohne einer kritischen Auseinandersetzung mit den Wissensquellen zu bedürfen (Bendixen & Rule, 2004; Hofer & Pintrich, 1997; Muis et al., 2006). In der sophistizierteren Ausprägung wird Wissen als vorläufig, veränderlich, komplex miteinander verbunden, persönlich in Interaktion mit dem sozialen Umfeld konstruiert und als kritisch zu hinterfragen und zu rechtfertigen betrachtet (Bendixen & Rule, 2004; Hofer & Pintrich, 1997; Muis et al., 2006). Die individuellen epistemologischen Überzeugungen werden bei allen Lernprozessen (Hofer, 2004a) und zwar in Abhängigkeit vom jeweiligen soziokulturellen, akademischen und instruktionalen Kontext aktiviert (Hofer, 2005, 2006b; Muis et al., 2006). Sie operieren sowohl auf kognitiver als auch metakognitiver Ebene (Bendixen & Rule, 2004; Hofer, 2004a, 2005, 2016; Muis et al., 2006), da die Metakognition auf die Kognition zugreift und diese wiederum speist (Veenman et al., 2006).

Ein sophistizierteres epistemologisches Überzeugungssystem wirkt sich dabei förderlich auf multiple lernrelevante Faktoren und demnach förderlich auf den Lehrprozess eines Indiviuums aus, während weniger fortgeschrittene Überzeugungen eher lernhemmend sind (Hofer, 2001; Schommer-Aikins, 2002). Sophistizierte epistemologische Überzeugungen sind lernförderlich im Hinblick auf multiple lernrelevante Faktoren, wie Lernemotionen (z.B. Muis et al., 2015), intrinsische Motivation (z.B. Buehl & Alexander, 2005; Hofer, 1999), Selbstkonzept (z.B. Urhahne & Harms, 2006), Fachinteresse, Selbststeuerungsfähigkeit (z.B. Bromme, 2010), Kontroll- und Selbstwirksamkeitsüberzeugungen (z.B. Muis & Foy, 2007; Paulsen & Feldman, 1999) und unterstützen den Lerntransfer (Jacobson & Spiro, 1995), die Informationssuche (z.B. Mason et al., 2010b; Pieschl et al., 2013), das Textverstehen (z.B.

Bråten & Strømsø, 2004, 2006; Bråten et al., 2008; Schommer-Aikins, 2004) und den *Conceptual Change* (Franco et al., 2012). Sie führen insbesondere zu förderlichem Lernverhalten, indem passendere Lernstrategien eingesetzt werden (Hofer, 2001; Schommer-Aikins, 2002). Sophistizierte epistemologische Überzeugungen gehen mit dem Einsatz tiefenorientierter Lernstrategien einher, aber hängen negativ mit der oberflächenorientierten Wiederholungsstrategie zusammen. Unsophistiziertere epistemologische Überzeugungen sind dagegen negativ mit tiefenorientierten Lernstrategien assoziiert und positiv mit der oberflächenorientierten Wiederholungsstrategie (z.B. Paulsen & Feldman, 2007; Schworm & Gruber, 2017).

Die individuellen wissensbezogenen Überzeugungen beeinflussen nicht nur den Lernprozess sondern auch den Lehrprozess (Hofer, 2001). Die epistemologischen Überzeugungen von Lehrerinnen und Lehrern wirken auf ihr unterrichtliches Handeln (Lehrtätigkeit, Lehrstrategien) und ihre Erwartungen an die Schülerinnen und Schüler (z.B. Brickhouse, 1989; Brownlee et al., 2011; Buehl & Fives, 2009; Hammrich, 1998; Hashweh, 1996; Muis & Franco, 2010; Tsai, 2002, 2006) sowie indirekt über das unterrichtliche Handeln auch auf die epistemologischen Überzeugungen der Schülerinnen und Schüler (z.B. Hofer, 1999; Muis & Franco, 2010; Schoenfeld, 1988; Stodolsky, 1985).

Aus dem vielfältigen Einfluss der wissensbezogenen Überzeugungen auf den Lernprozess geht die besondere Relevanz eines sophistizierten Überzeugungssystems für Lehramtsstudierende hervor. Die Lehrerbildung umfasst in Deutschland das Studium, Praktika und das Referendariat. Angehende Lehrkräfte müssen kontinuierlich selbstgesteuert lernen, und zwar sowohl in den verschiedenen Wissensdomänen ihrer Unterrichtsfächer als auch im erziehungswissenschaftlichen Wissensbereich, um ihr professionelles Handeln weiterzuentwickeln und das Lernen von Schülerinnen und Schülern erfolgreich fördern zu können (Buehl & Fives, 2016). Wie erfolgreich die eigenen Lernprozesse sind, wird allerdings von den individuellen Überzeugungen zum Wissen und zum Wissensprozess mit beeinflusst. Wissensbezogene Überzeugungen wirken als Linse, durch die Lernende Lernaufgaben wahrnehmen und Lerninhalte erfassen. Sie dienen der erfolgreichen Anpassung des Lernverhaltens – im Sinne einer förderlichen Lernstrategienauswahl – an die wahrgenommenen Anforderungen und Rahmenbedigungen der Aufgabe (Bromme, 2010; Pieschl et al., 2013). Die epistemologischen Überzeugungen filtern, wie Lehramtsstudierende das Lernen in ihren universitären Veranstaltungen – die ja deren Studium zu Lehrkräften dienen – erleben und inwiefern sie sinnvolle, tiefenorientierte Lernansätze betreiben (Many et al., 2002; Muis, 2004; Peng & Fitzgerald, 2006; Yadav & Koehler, 2007). Sie wirken sich folglich darauf aus, wie tief ihr in den universitären Veranstaltungen (*teacher education courses*) erworbenes Verständnis geht und beeinflussen außerdem ihre Entscheidungsfindung, Planung, Durchführung und Bewertung der praktischen Anwendung des Gelernten (Buehl & Fives, 2016; Yadav & Koehler, 2007). Auch die selbstständige Recherche und das selbstständige Lesen

und Erarbeiten von Texten, die wichtige Aspekte eines Studiums darstellen, sind vom epistemologischen Überzeugungssystem abhängig. Sophistizierte epistemologische Überzeugungen beeinflussen das Verhalten von Lernenden bei der selbstregulierten Informationssuche und -bewertung positiv (Hofer, 2004b; Mason et al., 2010b; Pieschl et al., 2013; Whitmire, 2004). Je fortgeschrittener das epistemologische Überzeugungssystem ist, desto eher wenden Studierende viel Zeit und Sorgfalt bei der Planung der Informationssuche auf, nutzen multiple Quellen, differenzieren zwischen der Herkunft von Inhalten und vergleichen kritisch zwischen verschiedenen Inhalten. Sophistizierte epistemologische Überzeugungen wirken sich förderlich auf die Textverarbeitung und das Textverstehen aus (z.B. Bråten & Strømsø, 2004; Kardash & Noel, 2000; Schommer et al., 1992; Schommer & Walker, 1995; Schommer-Aikins, 2004; Schraw et al., 2002). Ein sophistiziertes epistemologisches Überzeugungssystem wirkt sich lernförderlich auf alle Phasen des selbstregulierten Lernens aus (z.B. Bromme, 2010; Bromme et al., 2014; Stahl & Bromme, 2007), während sich ein weniger sophistiziertes Überzeugungssystem eher lernhemmend auswirkt (Hofer, 2001; Schommer-Aikins, 2002).

Schon bei angehenden Lehrerinnen und Lehrern stehen die individuellen wissensbezogenen Überzeugungen in starker Verbindung mit der unterrichtlichen Zielsetzung und den im Unterricht umgesetzten Handlungen (z.B. im Rahmen von Praktika oder des Referendariats) (z.B. Kang, 2008; Yadav & Koehler, 2007) und sophistizierte epistemologische Überzeugungen sind mit einer stärker konstruktivistisch ausgeprägten Didaktik verbunden (Brownlee et al., 2016). Angehende Lehrkräfte mit sophistizierteren epistemologischen Überzeugungen nutzen eher pädagogische Praktiken, die die Kinder ermutigen, unabhängig zu denken und Wissen zu konstruieren, als solche mit weniger sophistizierten epistemologischen Überzeugungen (Brownlee et al., 2016).

Für Lehramtsstudierende sind die eigenen epistemologischen Überzeugungen demnach in zweifacher Hinsicht von besonderer Bedeutung, denn sie beeinflussen nicht nur das eigene Lernen im Studium sondern auch die spätere Konzeption von Lernkontexten für Schülerinnen und Schüler (Buehl & Fives, 2016, S. 248). Für sie ist es besonders wichtig, über ein sophistiziertes epistemologisches Überzeugungssystem zu verfügen,

> „that can influence the depth of understanding they take from teacher education courses, the development of deep and varied knowledge systems for teaching, and the use of effective, informed decision-making across a range of contexts“ (Brownlee et al., 2011, S. 303).

Da epistemologische Überzeugungen aber disziplinspezifisch sind (Muis et al., 2006), wäre es erstrebenswert, dass Lehramtsstudierende sowohl bezüglich der Wissensdomänen ihrer jeweiligen Unterrichtsfächer als auch bezüglich des erziehungswissenschaftlichen Wissens über sophistizierte Überzeugungsysteme verfügen. Das erziehungswissenschaftliche Überzeugungssytem ist insofern zentral, als dass es die pädagogischen und didaktischen Inhalte umfasst, die für den späteren Lehrberuf wichtig sind (wie z.B. Unterrichtsplanung, Klassenführung, Umgang mit Unterrichtsstörungen, etc.) und sich auf alle Unterrichtsfächer auswirkt. Die Förderung sophistizierter epistemologischer Überzeugungen – explizit auch zum erziehungswissenschaftlichen Wissen – sollte demnach in der Lehrerbildung ein

zentrales Ziel darstellen (vgl. Brownlee et al., 2011, 2016). Die Voraussetzung der zielgerichteten Förderung ist jedoch die Identifikation des Ist-Zustandes. Nur wenn das aktuell vorherrschende epistemologische Überzeugungssystem ermittelt werden kann, lässt sich feststellen, bei wem und in Bezug auf welche Überzeugungsdimensionen Förderungsbedarf vorliegt (vgl. Chai et al., 2006). Und nur wenn die epistemologischen Überzeugungen erhoben werden können, kann der Erfolg einer durchgeführten Förderungsmaßnahme ermittelt werden. Für die verschiedensten Fachdisziplinen wie beispielsweise Mathematik, Physik und Geschichte ist die quantitative Erhebung mittels Fragebogen bereits möglich, da es passende Instrumente gibt. Für die Überzeugungen zum erziehungswissenschaftlichen Wissen war dies bisher aber nicht der Fall. Unter den zahlreichen quantitativen Erhebungsinstrumenten gab es bisher keines, das sich explizit für die Erfassung der Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen und Wissensprozess eignete.

Diese Arbeit hat sich demnach zum Ziel gesetzt, das StEB (students' epistemological beliefs) Inventar zu entwickeln, das die mehrdimensionale und disziplinspezifische Erfassung epistemologischer Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen aus Studientexten erlaubt.

Die Eingrenzung auf erziehungswissenschaftliches Wissen aus Studientexten erfolgte, da die epistemologischen Überzeugungen im Kontext aktiviert werden (Hofer, 2005, 2006b; Muis et al., 2006). Der akademische sowie der instruktionale Kontext beeinflussen, welche epistemologischen Überzeugungen in den Lehramtsstudierenden aktiviert werden und demnach auch erhoben werden können. Um möglichst sicherzustellen, dass die befragten Lehramtsstudierenden die „richtigen" epistemologischen Überzeugungen aktivieren, d.h. diejenigen, die hier von Interesse sind, wurde der instruktionale Kontext in dieser Instrumententwicklung auf Studientexte eingegrenzt. Dieser Aspekt wurde gewählt, da das selbstständige Lesen und Lernen aus Texten im Sinne des selbstgesteuerten Lernens ein wichtiger Aspekt des erziehungswissenschaftlichen Lehramtsstudiums ist und die disziplinspezifischen epistemologischen Überzeugungen die Textverarbeitung und das Textverstehen bekanntermaßen stark beeinflussen (z.B. Bråten & Strømsø, 2004; Kardash & Noel, 2000; Schommer et al., 1992; Schommer & Walker, 1995; Schommer-Aikins, 2004; Schraw et al., 2002).

Aus dieser Zielsetzung leitete sich die zugehörige Fragestellung ab, die danach fragt, ob sich die epistemologischen Überzeugungen von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen aus Studientexten in der Faktorstruktur des StEB konform zur vorgestellten theoretischen Synthese abbilden lassen. Der angestrebte Nutzen und spätere intendierte Einsatz des StEB Inventars wurde in zwei Testwertinterpretationen spezifiziert. Aus den Testwertinterpretationen und auf Basis der Theorie wurden im Anschluss acht zu prüfende Grundannahmen abgeleitet, die später der Validierung der Testwertinterpretationen dienen.

Im methodischen Hintergrund dieser Arbeit wurde zunächst über die grundsätzliche Logik der quantitativen Erfassung von Konstrukten im Rahmen von psychologischen Tests

zur klassischen Testtheorie (KTT) hingeleitet. Die Grundanahmen der KTT sowie die verschiedenen Messmodelle wurden knapp aufgeführt, bevor die Testgütekriterien Objektivität, Reliabilität und Validität beschrieben wurden, die dieser Testentwicklung zugrundeliegen. Im Zusammenhang mit der Validität wurde zunächst auf das alte Validitätsverständnis eingegangen, nach dem verschiedene Validitätsarten, wie die Konstruktvalidität, Inhaltsvalidität etc. unterschieden werden. Nach dem neuen, argumentationsbasierten Validierungsansatz (vgl. AERA et al., 2014; Hartig et al., im Druck), dem diese Arbeit folgt, werden diese Validitätsarten jedoch nicht mehr unterschieden und statt von der Validierung eines Tests zu sprechen, werden nun die angestrebten Testwertinterpretationen validiert. Demensprechend wurden die einzelnen Schritte des argumentationsbasierten Validierungsansatzes vorgestellt. Dabei wurde auf die Spezifikation von Testwertinterpretationen, das Aufstellen empirisch prüfbarer Grundannahmen, das Sammeln empirischer Evidenzen für und gegen die Grundannahmen sowie die zusammenfassende Beurteilung der Evidenzen eingegangen. Im Anschluss wurde der idealtypische Ablauf einer Instrumententwicklung nach der KKT unter Einsatz der faktorenanalytischen Verfahren skizziert, indem zunächst die Teilschritte der exploratorischen Faktorenanalyse und anschließend die der konfirmatorischen Faktorenanalyse aufgezeigt wurden.

Unter Rückgriff auf die Konstrukteingrenzung und Definition des Messgegenstandes wurde ein erster StEB Testentwurf entwickelt. Das umfasste die Bestimmung des Itemformats und die Ableitung von zunächst fünf Dimensionen epistemologischer Überzeugungen aus der Theorie. Zu diesen fünf Dimensionen, den Überzeugungen zur *Quelle*, *Rechtfertigung*, *Struktur*, *Sicherheit* und *Entwicklung des Wissens* wurde ein Pool von 126 Items entwickelt. Bevor diese Items jedoch einer ersten Stichprobe von Lehramtsstudierenden vorgelegt werden konnten, wurde deren Inhalt und Zuordnung zu den Überzeugungsdimensionen zunächst geprüft. Die Expertise von zehn Expertinnen und Experten aus den Bereichen Pädagogische Psychologie, Erziehungswissenschaft, Psychologie und Pädagogik von fünf deutschen Universitäten und des ZPID-Leibniz Instituts in Trier wurde genutzt, um die Itemauswahl, -inhalt und -formulierung zu optimieren sowie die interne Teststruktur zu prüfen. Die Expertinnen und Experten gaben einerseits ihre Einschätzung über die Eignung der Items in Bezug auf die jeweils intendierte Skala auf einer elfstufigen Skala von 1 (*Item unpassend*) bis 11 (*Item passend*) an. Darüber hinaus lieferten sie wertvolles Feedback zum Iteminhalt. Als Maß für die zusammengefasste Expertenmeinung dienten die Itemmediane, die ins Verhältnis zu der mitteren Einschätzung der zugehörigen Überzeugungsskala gesetzt wurden, um ungeeignete Items zu identifizieren. Auf Basis des Expertenfeedbacks und der -einschätzungen wurden die Items überarbeitet und die Itemauswahl optimiert. Es resultierte ein 98-Item StEB Testentwurf.

Die empirische Prüfung des 98-Item StEB Testentwurfs, im Rahmen der exploratorischen Modellbildung, umfasste den Einsatz an einer Stichprobe von zunächst 221 Lehramtsstudierenden zweier deutscher Universitäten. Nach der Bereinigung der Daten verblieben 217 Datensätze, von denen 187 bezüglich der Überzeugungsitems vollständig waren. Die Analyse fehlender Werte ergab eine insgesamt geringe Anzahl fehlender Werte und darüber hinaus

auch, dass den fehlenden Werten keine Systematik zugrunde liegt (MCAR), was den Einsatz der *arithmetic mean imputation* als Methode zum Umgang mit den fehlenden Werten erlaubt. Die fehlenden StEB-Itemwerte wurden durch den Stichprobenmittelwert der jeweils betroffenen Variable ersetzt. Neun Überzeugungsitems wurden daraufhin von den weiteren Analysen ausgeschlossen, da sich im Rahmen der Prüfung der Verteilungen herausstellte, dass die Befragten in ihrer Itembeantwortung nicht die volle Breite des Antwortspektrums abdeckten. Der Analyse der Itemverteilungen schloss sich die Prüfung der Voraussetzungen für den Einsatz der EFA an. Hier wurde analysiert, ob die Stichprobengröße ausreichend ist und mittels KMO-Kriterium und itemspezifischem MSA untersucht, ob sich die Itemkorrelationsmatrix für die EFA eignet. Dies hatte den Ausschluss von drei weiteren Items aufgrund zu geringer MSA-Werte zur Folge. Im Anschluss wurde die Hauptachsenanalyse als geeignetes verteilungsfreies Extraktionsverfahren gewählt, da der Mardia-Test ergab, dass für die Überzeugungsitems keine multivariate Normalverteilung angenommen werden kann und sich dieses Verfahren auch speziell für Ratingitems eignet. Zur anschließenden Bestimmung der geeigneten Faktorenanzahl für die EFA wurden neben den inhaltlich-theoretischen Überlegungen die Ergebnisse der Parallelanalyse, des Scree-Tests, der MAP-Methode sowie der OC- und AF-Methode berücksichtigt. Ein Vergleich der Ergebnisse der fünf berechneten EFAs (mit obliquer Rotation) ergab, dass die Extraktion von fünf Faktoren zu einem inhaltlich-theoretisch interpretierbaren Ergebnis führte. Die resultierende Mustermatrix wurde bezüglich der Höhe der Ladungen und Kommunalitäten untersucht, die neben dem Iteminhalt eine bedeutende Rolle bei der Itemreduktion spielten. Nach der Itemreduktion schloss sich eine erneute faktorenanalytische Prüfung der Faktorstruktur an, die ergab, dass die Faktorstruktur der verbliebenen 34-Item StEB-Version stabil bleibt. Für die 34-Item StEB-Version wurden im Rahmen einer Item- und Skalenanalyse die psychometrischen Itemschwierigkeiten und Itemtrennschärfen sowie die Skalenhomogenitäten und -reliabilitäten bestimmt. Dem schloss sich die Diskussion der ersten Studie an, die einen Rückbezug zur Fragestellung und der Theorie umfasste. Die aus der ersten Studie resultierende 34-Item StEB Inventar Version wurde abschließend vorgestellt. Sie verfügte über die vier Dimensionen Überzeugungen zur *Absolutheit des Wissens*, *Einfachheit des Wissens*, *Multimodalität des Wissenserwerbs* und *Entwicklung des Wissens*.

Das auf Basis der ersten Stichprobe modifizierte StEB Inventar wurde, im Sinne des idealtypischen Ablaufs einer Instrumententwicklung nach der KTT, im Rahmen der zweiten Studie mit neuen, größeren Stichprobe geprüft (vgl. Brown, 2015; Bühner, 2011; Fabrigar et al., 1999; R. B. Kline, 2016, S. 198). Zum Einsatz kam dazu der modellgenerierende Ansatz der konfirmatorischen Faktorenanalyse (CFA) (vgl. Jöreskog, 1993), da die CFA in dieser Art als ein Instrument zur Konstruktion und Evaluation von Skalen eingesetzt werden kann. Die 34-Item StEB-Version wurde gemeinsam mit dem Inventar zur Erfassung von Lernstrategien im Studium (LIST-R, revidierte Version, Wild, 2011) an einer Stichprobe von zunächst 1051 Lehramtsstudierenden sechs deutscher Hochschulen eingesetzt. Nach der Bereinigung der Daten und dem Ausschluss von acht Ausreißern verblieben die Datensätze von schlussendlich $N = 915$ Lehramtsstudierenden. Die folgende Analyse der fehlenden StEB-Werte lehnte

die MCAR Annahme ab. Aufgrund der starken Hinweise auf MAR wurde als geeignetes Verfahren zum Umgang mit den fehlenden StEB-Werten FIML gewählt. MLR wurde nach der Prüfung der CFA-Voraussetzungen bezüglich Skalenniveau, Verteilung und Stichprobengröße der ML-Methode als die angemessene Schätzmethode vorgezogen, da sie keine multivariat normalverteilten Daten benötigt und die Tatsache berücksichtigt, dass Ratingitems vorliegen. Die Spezifikation des ersten CFA-Gesamtmodells wurde auf Basis der aus der ersten Studie resultierenden Faktorstruktur vorgenommen. Sowohl die Parameterschätzungen des standardisierten und unstandardisierten Gesamtmodells als auch die Modifikationindizes und die Residualkovarianzmatrix der Items wurden genau untersucht. Um den identifizierten Missspezifikationen nachzugehen, wurden im Anschluss die vier einfaktoriellen CFA-Modelle (ein Modell je Faktor) spezifiziert und geprüft sowie eine exploratorische Prüfung des Gesamtmodells mittels EFA vorgenommen. Die EFA erleichterte die Interpretation der durch die CFA implizierten notwendigen Modellmodifikationen, indem den Items die Zuordnung zu den Faktoren freigestellt wurde und sich so Unstimmigkeiten in der für das Gesamtmodell spezifizierten Item-Faktor-Zuordnung erkennen ließen. Wie sich in der EFA zeigte, konnte die aus der ersten Stichprobe resultierende Faktorstruktur des StEB Inventars weitgehend repliziert werden. Die Ergebnisse der CFAs zum vierfaktoriellen Gesamtmodell und zu den einfaktoriellen Modellen sowie der Vergleich dieser Ergebnisse mit denen der EFA implizierten verschiedene Modellanpassungen. Die Modellanpassungen bestanden jedoch nur – mit der Ausnahme eines Items, das den Faktor wechselte – im Entfernen ungeeigneter Items. Um die Modellanpassungen kreuzvalidieren zu können, d.h. zu prüfen, ob sie zu einer stabilen Modelllösung führen, wurde die gesamte Validierungsstichprobe randomisiert in zwei Teilstichproben geteilt. So wurden Schritt für Schritt elf Modellrevisionen erst an Teilstichprobe 1 und danach an Teilstichprobe 2 vorgenommen. Für jeden Schritt wurden die Veränderungen in den Parameterschätzungen, den Modifikationsindizes und der Modellpassung beider Modelle geprüft und verglichen. So konnte untersucht werden, ob umgesetzte Modelländerungen in beiden Stichproben zu vergleichbaren Veränderungen in der Modellpassung führten. Neun dieser Modifikationen wurden für das finale Modell beibehalten, da diese neun vorgenommenen Modelländerungen sich in beiden Stichprobenteilen gleich auswirkten. Aus Studie 2 resultiert auf diese Weise eine vierfaktorielle 26-Item Modelllösung. Sowohl für die einfaktoriellen Modelle der Gesamtstichprobe als auch die einfaktoriellen Modelle der beiden Teilstichproben konnte *exakter* Modell-Fit erreicht werden, wenn je nach Modell keine bis maximal drei korrelierte Fehlervarianzen (Messfehler) spezifiziert wurden. Da die vorliegenden Zusammenhänge zwischen den Messfehlern der einfaktoriellen Modelle und des vierfaktoriellen Modells der Gesamtstichprobe (sowie der einfaktoriellen Modelle der Teilstichproben) äußerst gering waren, wurde davon ausgegangen, dass sie nur aufgrund der großen Stichprobe signifikant wurden. Demnach wurde die vierfaktorielle 26-Item Modelllösung ohne korrelierte Messfehler als abschließendes Modell des StEB Inventars gewählt. Das abschließende StEB Inventar verfügte damit weiterhin über die vier Dimensionen Überzeugungen zur *Absolutheit des*

*Wissens*, *Einfachheit des Wissens*, *Multimodalität des Wissenserwerbs* und *Entwicklung des Wissens*. Nach einer Diskussion der Ergebnisse und kurzen Darstellung der resultierenden Skalen wurde das finale StEB Inventar einer Messmodellprüfung sowie Item- und Skalenanalyse unterzogen. Im Rahmen der Itemanalyse wurden die psychometrischen Itemschwierigkeiten und part-whole korrigierten Itemtrennschärfen bestimmt. Die Messmodellprüfung zeigte, dass für das StEB Inventar das tau-kongenerische Messmodell angenommen werden kann. Für die finalen vier StEB-Dimensionen ergaben sich Reliabilitäten von $\omega$ = .70 bis $\omega$ = .82 (bzw. $\alpha$ = .70 bis $\alpha$ = .82) und Skalenhomogenitäten von *MIC* = .25 bis *MIC* = .36. Faktorwerte als Testwerte des StEB Inventars zu bestimmen, wäre zwar hinsichtlich der Messmodellprüfung konsequent gewesen, würde aber für den späteren Einsatz des StEB bedeutende Nachteile mit sich bringen. Aus diesem Grund wurde beschlossen, dass für das StEB Inventar Skalenmittelwerte als Testwerte verwendet werden.

Zur Validierung der Testwertinterpretationen des finalen 26-Item StEB Inventars nach dem argumentationsbasierten Ansatz wurden nun die sieben aus den Testwertinterpretationen abgeleiteten Grundannahmen geprüft. Die Prüfung diente dazu, Evidenzen zu sammeln, die die Testwertinterpretationen stärken oder aber gegen sie sprechen, um schlussendlich eine Aussage darüber treffen zu können, ob die intendierten Testwertinterpretationen und folglich der geplante Einsatz des StEB zulässig sind. Als Evidenzquellen wurde die interne Teststruktur, der Testinhalt und die Beziehung der StEB-Testwerte zu anderen Variablen herangezogen. Es konnte einerseits gezeigt werden, dass die Testwerte des StEB auf die verschiedenen Überzeugungsausprägungen zurückzuführen sind (Grundannahme 1) und dass diese wiederum mit der theoretischen Konzeptualisierung nach Hofer und Pintrich (1997) vereinbar sind (Grundannahme 2). Darüber hinaus konnten die angenommenen Zusammenhänge der StEB-Faktoren untereinander sowohl gefunden, als auch argumentativ hergeleitet werden, dass sich diese bestehenden Zusammenhänge konstruktkonform verhalten (Grundannahme 3). Zur Prüfung der Grundannahme 4 wurden, einem korrelativen Ansatz folgend, die empirisch vorliegenden Beziehungen zwischen den StEB- und LIST-R-Testwerten mit dem zuvor aufgestellten nomologischen Netz abgeglichen. Für eine Prüfung mittels Strukturgleichungsmodellen stellte sich der LIST-R als ungeeignet heraus. Da sophistiziertere epistemologische Überzeugungen zum erziehungswissenschaftlichen Wissen hypothesenkonform mit tiefenorientierten Lernstrategien und unsophistiziertere Überzeugungen mit Oberflächenstrategien einhergingen, konnte auch Grundannahme 4 als bestätigt gelten. Die Grundannahmen 5 bis 7 postulierten Unterschiede in den Überzeugungsausprägungen der vier Dimensionen in Abhängigkeit der Hochschule, der Fachsemester und der fachlichen Ausrichtung in Kombination mit den Fachsemestern. Sie wurden mittels MANOVAs geprüft. Wie angenommen, zeigten sich signifikante Unterschiede in den wissensbezogenen Überzeugungen der Personengruppen unterschiedlicher Hochschulen (Grundannahme 5). Die Prüfung von Grundannahme 6 lieferte keine eindeutigen Ergebnisse, da nur sehr schwache Unterschiede gefunden wurden. Über Grundannahme 7 ließ sich keine Aussage treffen. Insgesamt zeigten sich für die gefundenen Unterschiede der MANOVAs nur sehr

geringe Effektstärken. In einer zusammenfassenden Beurteilung der gesammelten Ergebnisse des Validierungsprozesses wurde geschlussfolgert, dass die Evidenzen die angestrebten Testwertinterpretationen insgesamt stärken. Dies spricht dafür, dass die StEB-Testwerte in der intendierten Art und Weise interpretiert werden dürfen und das StEB Inventar entsprechend eingesetzt werden kann.

Das StEB Inventar eignet sich folglich dazu, die bei Lehramtsstudierenden aktuell vorherrschenden Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten mehrdimensional zu ermitteln. Es erlaubt darüber hinaus einen intraindividuellen und interindividuellen Vergleich der mehrdimensionalen Ausprägungen epistemologischer Überzeugungen zum erziehungswissenschaftlichen Wissen aus Studientexten.

Daraus ergeben sich forschungspragmatische und praktische Anwendungsmöglichkeiten für das StEB Inventar, die in Kapitel 9.5 (S. 338) aufgezeigt werden. Im folgenden Kapitel 9.2 wird zunächst jedoch der Bezug des StEB Inventars zu der entwicklungspsychologischen Sichtweise auf epistemologische Überzeugungen dargestellt, bevor im Anschluss in Kapitel 9.3 die normative Sichtweise auf epistemologische Überzeugungen diskutiert wird.

## 9.2 Bezug des StEB zu den Entwicklungsstufen epistemologischer Überzeugungen

In Kapitel 2.9.3 (ab S. 122) wurde bereits auf die Problematik eingegangen, die bei manchen bestehenden Fragebögen (z.B. Hofer, 2000) zur Erfassung epistemologischer Überzeugungen bezüglich der Interpretierbarkeit der mittleren Ausprägung vorliegt. Das betrifft Fragebögen, bei denen die epistemologischen Überzeugungen als Selbstauskünfte mittels Likertskalen erfasst werden und denen gleichzeitig die Annahme zugrundeliegt, sie seien mit der entwicklungspsychologischen Sichtweise auf epistemologische Überzeugungen vereinbar (Mason, 2016). Mason (2016) kritisiert, dass nicht angenommen werden kann, dass die Itembeantwortung bei Likertitems die drei epistemologischen Ausprägungen *Absolutismus*, *Relativismus* und *Evaluatismus* linear abbilden können. Die Interpretation der mittleren Kategorie der Likertskala im Sinne einer relativistischen Überzeugungsausprägung sei nicht zulässig. Kritisiert wurde nicht nur die Interpretierbarkeit der mittleren Ausprägung, sondern es wurde auch angezweifelt, ob nicht der rechte Pol generell eher einer relativistischen Sichweise auf Wissen entspricht und der linke Pol einer absolutistischen (Hofer, 2016). Somit würde der rechte Pol zwar immer noch die sophistiziertere Überzeugung abbilden, allerdings wäre die höchste Entwicklungsstufe, der Evaluatismus, gar nicht mehr abgebildet. Beide Kritikpunkte erscheinen inhaltlich sinnvoll. Betrachtet man beispielsweise die Überzeugungen zur Sicherheit des Wissens, würde die Annahme, dass Wissen komplett sicher ist, dem wenig sophistizierten Absolutisten entsprechen. Die gegenteilige Annahme, dass Wissen komplett unsicher ist, bildet allerdings keine evaluatistische, sondern höchstens eine relativistische Stufe der epistemologischen Entwicklung ab. Somit würde ein entsprechendes Item nur zwei der drei Entwicklungsschritte abbilden. Ein Kritikpunkt bezüglich der bisherigen Erhebungsinstrumente, die Likertskalen einsetzen, ist demnach, dass diese einerseits einen linearen Entwicklungsverlauf annehmen und andererseits mit dem entwicklungspsychologischen Ansatz epistemologischer Überzeugungen vereinbar sein wollen, obwohl sie die Merkmalsausprägung Evaluatismus inhaltlich möglicherweise gar nicht erfassen können, sondern nur die Extreme Absolutismus und Relativismus abdecken (Hofer, 2016) oder aber nur die beiden Extreme Absolutismus und Evaluatismus abbilden (Mason, 2016).

Aufgrund dieser Problematik wurde für das StEB Inventar von Anfang an nicht davon ausgegangen, dass es auf Itemebene mit der entwicklungspsychologischen Sichtweise vereinbar ist. Es wird demnach *nicht* davon ausgegangen, dass die weniger sophistizierte Ausprägung immer der absolutistischen Überzeugung zu Wissen und die sophistizierte immer der evaluatistischen Überzeugungen entspricht, während eine multiplizistische Sichtweise auf Wissen möglichwerweise irgendwo in der Mitte anzusiedeln ist. Vielmehr wurde für jeden Entwicklungsschritt analysiert, was die hohen bzw. niedrigen Ausprägungen der Überzeugungsdimensionen inhaltlich bedeuten. In den Tabellen 5.2, 7.20 und 6.14 (S. 177, 261, 212) zur Item-Skalen-Zuordnung wurden die Ergebnisse festgehalten.

Wie sich jedoch bei Analyse der Überzeugungsausprägungen zeigt, lassen sich die Skalen des StEB teilweise mit der entwicklungspsychologischen Sichtweise (z.B. Kuhn, 1991;

Kuhn & Weinstock, 2002) in Einklang bringen. Die beiden Extreme der Dimension *Absolutheit des Wissens* lassen sich der absolutistischen und relativistischen Sichtweise auf Wissen zuordnen, während die Pole der Dimension *Multimodalität des Wissenserwerbs* eine absolutistische und evaluatistische Sichtweise abbilden. Für die anderen beiden StEB-Dimensionen, die *Entwicklung* und *Einfachheit des Wissens*, ist die Zuordnung nur für die unsophistizierte Ausprägung eindeutig absolutistisch, der jeweils sophistizierte Pol scheint eher eine Mischung aus relativistischer und evaluatistischer Sichweise auf Wissen widerzuspiegeln.

Für die Skala *Einfachheit des Wissens* entspricht eine Zustimmung zu den Items einer angenommenen *einfachen* Struktur des Wissens. Lehramtsstudierende mit dieser Überzeugung sind der Ansicht, dass erziehungswissenschaftliches Wissen aus Studientexten aus sicheren, diskreten Fakten besteht, die es zu wissen gilt. Würde man diese Überzeugungsausprägung in ein Verhältnis zur entwicklungspsychologischen Sichtweise setzen, würde man dieser Ausprägung die absolutistische Sichtweise auf Wissen zuordnen. Allerdings entspricht der Gegenpol nicht unbedingt einer evaluatistischen Sichtweise, sondern vielleicht noch eher einer Mischung aus Relativismus und Evaluatismus. Lehramtsstudierende, die die Items ablehnen, sind der Überzeugung, dass erziehungswissenschaftliches Wissen aus Studientexten aus unsicheren, untereinander verbundenen Konzepten besteht, d.h. eine *komplexe* Struktur aufweist. Die untereinander verbundenen Konzepte wären wahrscheinlich einer evaluatistischen Sichweise auf Wissen zuzuordnen, während die angenommene Unsicherheit dieser komplexen Struktur nicht unbedingt eine evaluatistische Ansicht ausdrücken muss, sondern viel eher im Relativismus verortet werden kann. Die Überzeugung zur einfachen Wissensstruktur wird als unsophistiziert bezeichnet, während die Überzeugung zur komplexen Struktur als sophistizierte Ausprägung betrachtet wird.

Die Skala *Entwicklung des Wissens* verfügt über die Pole *statisch* und *entwickelnd*. Lehramtsstudierende, die den Items zustimmen, sind der Überzeugung, dass das erziehungswissenschaftliche Wissen, das in Studientexten dargestellt wird, sich über die Zeit verändert, wodurch eine neue Auseinandersetzung mit den Inhalten nötig wird. Lehramtsstudierende mit Überzeugungen des anderen Extrems glauben vielmehr, dass das erziehungswissenschaftliche Wissen aus Studientexten über die Zeit unverändert bleibt, weshalb keine erneute Auseinandersetzung mit den Inhalten erforderlich ist. Die entwickelnde Sichtweise auf Wissen wird als sophistiziert betrachtet, die statische als unsophistiziert. Auch bezüglich dieser Skala ist die Zuordnung der entwicklungspsychologischen Beschreibung für den sophistizierten Pol nicht eindeutig. Die Ansicht, dass erziehungswissenschaftliches Wissen statisch ist, wäre sicher dem Absolutismus zuzuordnen. Die Sichtweise, dass sich Wissen weiterentwickelt, könnte allerdings sowohl Relativismus oder auch Evaluatismus ausdrücken.

Für die Skala *Multimodalität des Wissens* entspricht die Zustimmung zu den Items der *multimodalen* Überzeugungsausprägung, die als sophistiziert betrachtet wird. Lehramtsstudierende, die zustimmend antworten, sind der Ansicht, dass der eigene erziehungswissenschaftliche Wissenserwerb einer vielfältigen, elaborierten Auseinandersetzung mit erziehungswissenschaftlichen Studientexten bedarf. Sie glauben, dass dieser Wissenserwerb auf das eigene Verstehen der Inhalte abzielt und die eigenständige Auseinandersetzung mit

verschiedenen Wissensquellen, Forschungsmeinungen, Sichtweisen, Informationen aus verschiedenen Studientexten erfordert. Die andere Überzeugungsausprägung, die sich in einer Ablehnung der Items ausdrückt, wird als *unimodal* bezeichnet. Lehramtsstudierende mit dieser wissensbezogenen Sichtweise sind der Überzeugung, dass eine einfache Auseinandersetzung mit den vorgegebenen erziehungswissenschaftlichen Studientexten für den erziehungswissenschaftlichen Wissenserwerb ausreicht. Sie sind der Ansicht, dass das einfache Lesen eines Studientextes als Auseinandersetzung ausreicht und weitere Wissensquellen, Forschungsmeinungen und Sichtweisen für den erziehungswissenschaftlichen Wissenserwerb nicht relevant sind. Dieser Pol wird als die unsophistizierte Überzeugungsausprägung beschrieben. Werden diese beiden Extreme dem entwicklungspsychologischen Ansatz zugeordnet, ließe sich die unsophistizierte Überzeugung darüber, dass reines Lesen als Auseinandersetzung genügt, der absolutistischen Sichweise zuordnen. Die sophistizierte Überzeugung zum eigenen Wissenserwerb kann der evaluatistischen Sichweise auf Wissen zugeordnet werden. Hier kommt es eindeutig auf die Qualität der Auseinandersetzung an, die auf das eigene Verstehen abzielt, weshalb der Relativismus diese Überzeugungsausprägung nicht angemessen beschreiben würde.

Bezüglich der Skala *Absolutheit des Wissens* entspricht eine Zustimmung zu den Items der Skala einer *absoluten* und eine Ablehnung einer *relativen* Sichtweise auf Wissen (siehe Tab. 7.20, S. 261). Lehramtsstudierende, die die Items stärker ablehnen, sind eher der Überzeugung, dass erziehungswissenschaftliche Studientexte eine unzuverlässige Quelle darstellen, deren Inhalte anzuzweifeln sind. Lehramtsstudierende, die den Items stärker zustimmen, sind der Überzeugung, dass erziehungswissenschaftliche Studientexte eine verlässliche Quelle für erziehungswissenschaftliches Wissen darstellen, deren Inhalte man ohne Hinterfragen glauben sollte. Setzt man diese Skala in ein Verhältnis zur entwicklungspsychologischen Sichtweise, fällt auf, dass sie in ihren Polen den absolutistischen und relativistischen Aspekt abdeckt, aber kein Anspruch darauf erhoben wird, dass diese Skala auch die evaluatistische Sichtweise auf Wissen erfassen kann. Im Kontext der vorherigen Ausführungen wurden Überzeugungen des relativen Pols als sophistizierter bezeichnet und Überzeugungen des absoluten Pols als unsophistizierter.

## 9.3 Normative Sichtweise auf epistemologische Überzeugungen

In diesem Zusammenhang stellt sich dann allerdings die Frage, ob die Einteilung in sophistizierte wissensbezogene Überzeugungen und unsophistizierte Überzeugungen angemessen ist. Ist es denn zulässig, Lehramtsstudierenden eine größere Sophistiziertheit in ihren Überzeugungen zu unterstellen, die im Sinne geringer Absolutheitsausprägungen alles anzweifeln und hinterfragen, was in erziehungswissenschaftlichen Studientexten steht? Oder könnte diese relative Sichtweise auf Wissen nicht auch dazu führen, dass die Studierenden selbst gut begründete und belegte Argumente erst gar nicht in Betracht ziehen, weil sie grundsätzlich zunächst einmal alles anzweifeln? Auch über die Qualität des Hinterfragens

und Anzweifelns werden hier keine Aussagen getroffen. Handelt es sich um ein begründetes, abwägendes Hinterfragen, oder um ein grundsätzliches Ablehnen ohne Berücksichtigung von Argumenten und Evidenzen? Im zweiten Fall entspräche die epistemologische Sichtweise der Überzeugungsausprägung, die Perry (1981) als Multiplizismus (Position 3 und 4) beschreibt (vgl. Abschnitt 2.2.1, ab S. 25). Auch wenn diese epistemologische Sichtweise auf Wissen nicht so fortgeschritten und lernförderlich ist, wie eine weiterentwickeltere, evaluatistische Überzeugung zu Wissen, stellt sie im Vergleich zur absoluten Ausprägung dennoch einen Entwicklungsfortschritt dar und kann somit tatsächlich als sophistiziertere Ausprägung der beiden beschrieben werden.

Diese normative Sichtweise auf die epistemologischen Überzeugungen, bei der eine Ausprägung als sophistiziert und fortgeschritten, die andere aber als *naiv* und weniger fortgeschritten beschrieben wird, wurde schon von beispielsweise Muis (2004), Bromme, Kienhues und Stahl (2008) und Ferguson und Bråten (2013) kritisiert. Muis (2004) schlug aufgrund der negativen Konnotation des Begriffs *naiv* vor, die Begrifflichkeiten förderlich (*availing*) vs. nicht förderlich (*non-availing*) statt *sophistiziert* vs. *naiv* zu verwenden. Nach diesem Vorschlag sollten epistemologische Überzeugungen, die mit besseren Lernergebnissen einhergehen als *förderlich* (*availing*) bezeichnet werden, während solche, die keinen oder einen negativen Einfluss auf das Lernen haben, als *nicht förderlich* (*non-availing*) bezeichnet werden sollen. Bromme et al. (2008) und Kienhues et al. (2011) sprechen sich dafür aus, dass vom Kontext abhängig ist, welche epistemologischen Überzeugungen förderlich sind oder eben in diesem Kontext als fortgeschritten angesehen werden können. In diesem Sinne würden sophistizierte epistemologische Überzeugungen einer Flexibilität im epistemologischen Denken entsprechen, d.h. der Fähigkeit oder Neigung je nach Kontext verschiedene epistemologische Standards und Praktiken anzuwenden (Brownlee, Ferguson & Ryan, 2017, S. 3).

Es sind durchaus Lernkontexte denkbar, in denen die Überzeugungen, dass Wissen aus sicheren Fakten besteht, sich vorteilhaft für Lernende auswirken können. Zielt die Leistungsermittlung einer universitären Lehrveranstaltung zu Semesterende auf das reine Abfragen deklarativen Wissens ab, wird eine Lehramtsstudierende, die aufgrund ihrer absoluten Überzeugungen beim Lernen Wiederholungsstrategien anwendet und alle Veranstaltungsinhalte auswendig lernt, möglicherweise eine bessere Note erzielen, als eine andere Lehramtsstudierende, deren Lernen aufgrund ihrer wissensbezogenen Überzeugungen auf das Verstehen ausgerichtet ist.

Die Frage ist allerdings, ob eine Wertung der Überzeugungssysteme nicht trotzdem zulässig ist. Schließlich führt das eine Überzeugungssystem eher zu Lernen, das auf Verstehen abzielt, während das andere eher zu Auswendiglernen führt. Nur aufgrund der ungünstigen Lernumgebung stellt sich in diesem Beispiel die *naivere* Überzeugungsausprägung als kurzfristig günstiger heraus, und das auch nur für die Leistungsbewertung. Die förderlichen Effekte der *sophistizierteren* Überzeugungen bleiben bestehen, nämlich die Tatsache, dass die zweite Lehramtsstudierende die Inhalte verstanden hat und später besser wird anwenden können. Sophistiziertere Überzeugungen gehen mit dem Einsatz oder der Neigung

zu tiefenorientierten Lernstrategien einher (z.B. Dahl et al., 2005; Hofer, 2001; Schommer-Aikins, 2002). Auch in der Studie 2 dieser Arbeit wurden Zusammenhänge zwischen den sophistizierten Überzeugungsausprägungen und den tiefenorientierten Lernstrategien gezeigt (siehe Tab. 8.7, S. 290). Zwischen den Überzeugungsskalen und den Wiederholungsstrategien zeigen sich die angenommenen Zusammenhänge häufiger jedoch nicht (z.B. Dahl et al., 2005; Phan, 2006; Urhahne & Hopf, 2004). Auch in Studie 2 dieser Arbeit (siehe Tab. 8.7, S. 290) zeigten sich die angenommenen negativen Zusammenhänge zwischen den Überzeugungen zur Multimodalität des Wissenserwerbs und der Entwicklung des Wissens mit den Wiederholungsstrategien nicht. Diese Tatsache könnte ein Hinweis dafür sein, dass Personen mit sophistiziertem Überzeugungssystem in der Lage sind, die Anforderungen der Lernsituation zu durchschauen und ihr Lernverhalten entsprechend anzupassen (vgl. auch Bromme, 2010; Pieschl et al., 2013). Die Studierende mit den sophistizierteren Überzeugungen mag also durchaus in der Lage sein, die Anforderungen der Lernsituation richtig einzuschätzen und zu erkennen, dass in diesem Fall reines Auswendiglernen gefragt ist und das Verstehen zumindest für die Klausur nicht erforderlich ist. Diese Studierende könnte sich demnach trotz ihres sophistizierten Überzeugungssystems dazu entscheiden, die Inhalte zunächst einmal auswendig zu lernen. Somit scheint die Bezeichnung des Überzeugungssystems dieser Studierenden als sophistiziert oder fortgeschritten wieder berechtigt, da mit diesem System die Fähigkeit einhergeht, sowohl komplexe als auch einfache Lernsituationen richtig einzuschätzen und das Verhalten entsprechend anzupassen. Diese Anpassungsfähigkeit konnten Bromme (2010) und Pieschl et al. (2013) bereits zeigen. Eine Person mit unsophistiziertem Überzeugungssystem, also beispielsweise der Überzeugung, dass Wissen feststehend ist und aus unverbundenen Fakten besteht, ist dagegen nur fähig, auf einfache Aufgaben (z.B. Vokabeln lernen) angemessen zu reagieren. Das erklärt dann auch, wieso die negativen Zusammenhänge der Überzeugungsdimensionen mit den Wiederholungsstrategien nicht unbedingt gefunden werden. Wenn eine Lernsituationen derart ist, dass das Wiederholen eine geeignete Strategie darstellt, wird eine Person mit sophistiziertem Überzeugungssystem dies erkennen, ihr Handeln anpassen und demnach Wiederholungsstrategien einsetzen, während eine Person mit unsophistiziertem Überzeugungssystem sowieso Wiederholungsstrategien einsetzt, weil sie generell dazu neigt.

Um eine begründete Entscheidung zu treffen, welche epistemologische Überzeugungsausprägung für Lehramtsstudierende als fortgeschrittener gelten sollte, lassen sich auch die Bildungsziele der Lehrerbildung heranziehen. Aus den Zielen der Lehrerbildung lässt sich ableiten, welche Überzeugungen zum erziehungswissenschaftlichen Wissen als die angestrebten gelten können. Dabei handelt es sich nämlich um die wissensbezogenen Überzeugungen, die den Erwerb der Kompetenzen begünstigen, auf die die Lehrerbildung abzielt und die im Beruf zum gewünschten pädagogischen Verhalten führen. Es gibt derzeit schon Untersuchungen, die für das System von wissensbezogenen Überzeugungen, das in dieser Arbeit als sophistiziert bezeichnet wurde, günstige Einflüsse auf das unterrichtliche Handeln nachgewiesen haben (siehe Abschnitt 2.8.3, S. 110). Aus diesen Gründen wurde in dieser Arbeit eine Unterscheidung in naive, wenig sophistizierte und sophistizierte, fortgeschrittene epistemologische Überzeugungen vorgenommen.

## 9.4 Einschränkungen

Bezüglich der Auswahl der Probanden ist festzuhalten, dass keine repräsentative Stichprobe vorliegt, da die Auswahl der Lehramtsstudierenden nicht randomisiert erfolgte. Gleichzeitig wurde jedoch insgesamt eine große Stichprobe erreicht und es wurden Lehramtsstudierende von sechs verschiedenen deutschen Universitäten befragt. Hier ist anzumerken, dass die Anzahl der erreichten Lehramtsstudierenden über die Universitäten hinweg stark variiert (siehe Tab. 7.2, S. 222). In Passau wurden die meisten Probanden erreicht, während beispielsweise in Potsdam nur 14 Studierende befragt wurden. Über den Akquiseprozess (siehe Kapitel 7.1, S. 217) hat die Autorin versucht, bei der Auswahl der Studierenden der anderen Hochschulen (MLU Halle, PH Freiburg, Uni Regensburg, Uni Bremen, Uni Potsdam) einen Selection-Bias zu verhindern, indem die Dozierenden der anderen Universitäten gebeten wurden, den Fragebogen in ihren Seminaren einzusetzen. So sollte sichergestellt werden, dass alle Studierenden eines Seminars teilnehmen und nicht nur die beispielsweise besonders motivierten oder interessierten Studierenden. Sowohl telefonisch als auch schriftlich wurde der Erhebungsprozess mit den Dozierenden abgesprochen. Der Ablauf der Erhebung konnte jedoch nur an der PH Freiburg und der Uni Passau kontrolliert werden. Die Dozierenden berichteten allerdings, den Papier-Fragebogen wie besprochen in den Veranstaltungen eingesetzt zu haben. Wie die Studienstandorte sind auch die unterschiedlichen Schularten unter den Probanden unterschiedlich häufig vertreten (siehe Tab. 7.3, S. 222). Hier besteht wieder eine Verbindung zur Akquise. Um möglichst viele Versuchspersonen zu erreichen, wurden häufiger Einführungsvorlesungen für die Erhebung ausgewählt. Gleichzeitig wurde die Befragung zum Ende des ersten Semesters durchgeführt, um sicherzustellen, dass die Studierenden überhaupt Gelegenheit hatten, erziehungswissenschaftliche Bildungserfahrungen gemacht zu haben. In diesen Einführungsvorlesungen sind die Schularten nicht gleich häufig vertreten. An der Uni Passau beispielsweise waren viel mehr Grundschullehramtsstudierende in der Einführungsvorlesung vertreten, da die Nachfrage nach Studienplätzen in diesem Bereich höher ist[1]. Die Grundschullehramtsstudierenden sind mit 352 Personen am stärksten vertreten. Die Schulart Hauptschule ist kaum vertreten, da es sich dabei um ein eher auslaufendes Schulmodell handelt, das beispielsweise in Bayern von der Mittelschule abglöst wird. Es ist jedoch trotzdem gelungen auch von den meisten anderen Schularten (Realschule, Mittelschule, Gymnasium und Sekundarschule) mindestens 100 Probanden oder mehr zu erreichen (siehe Tab. 7.3, S. 222).

Ein Selection-Bias ist bei der Onlinebefragung nicht auszuschließen. Zwar erhielten nur die Teilnehmerinnen und Teilnehmer vorher spezifisch ausgewählter Veranstaltungen Zugriff auf die Onlinebefragung, allerdings war der Rücklauf sehr gering. Die Dozierenden der entsprechenden Onlineveranstaltungen wurden gebeten, Werbung für die Teilnahme an der StEB-Befragung in den Onlinekursen einzustellen und die Studierenden zur Teilnahme aufzufordern und später nochmal zu erinnern. Inwiefern dies umgesetzt wurde, konnte von der Autorin nicht geprüft werden. Es ist also beispielsweise möglich, dass unter den

[1] Aus diesem Grund galt als Zulassungsbeschränkung in Passau im Wintersemester 2016/2017 ein NC der Note 2,7.

Studierenden, die an der Onlinebefragung teilnahmen und deren Daten verwertet werden konnten (weil sie hinreichend vollständig waren), solche Studierende stärker vertreten sind, die beispielsweise über sophistiziertere Überzeugungen und/oder über ein höheres erziehungswissenschaftliches Interesse verfügen und/oder erziehungswissenschaftlichem Wissen eine größere Bedeutung beimessen.

Die Ergebnisse der Unterschiedsprüfung mittels MANOVA sollten differenziert betrachtet werden. Sie lassen sich eher eingeschränkt zur Stützung der Gültigkeit der Testwerte heranziehen, da sie nur geringe Effekte erzielten. Insgesamt scheinen sich die Lehramtsstudierenden der verschiedenen Standorte ziemlich in ihren Überzeugungsausprägungen zu ähneln. Das mag einerseits daran liegen, dass die Hälfte der Probanden Erstsemester ($n$ = 455) sind, d.h. Personen mit nur einem Semester akademischer Erfahrung in der Disziplin Erziehungswissenschaft, die noch wenig Gelegenheit hatten, ihre erziehungswissenschaftlichen wissensbezogenen Überzeugungen weiterzuentwickeln. Andererseits zeigte aber auch die Unterschiedsprüfung bezüglich der Überzeugungsausprägungen der Gruppe der Studienanfänger (1. und 2. Fachsemester) und der Gruppe der fortgeschrittenen Studierenden (ab 3. Fachsemester), dass zwischen diesen beiden kaum Unterschiede in den Überzeugungen vorliegen. Hier zeigte sich nur ein Unterschied mit geringem Effekt in den Überzeugungen zur Einfachheit des Wissens. Das ist ein Hinweis darauf, dass die Studierenden – auch die höherer Semester – generell zu wenig erziehungswissenschaftliche Bildungserfahrungen gemacht haben, als dass sich diese auf die Entwicklung der erziehungswissenschaftlichen Überzeugungen hätte ausgewirken können. Nach dem TIDE Modell (Muis et al., 2006) entwickeln sich disziplinspezifische Überzeugungen, wie die Überzeugungen zum erziehungswissenschaftlichen Wissen, in akademischer Auseinandersetzung mit dem spezifischen Fach. Auch nach dem IM von Bendixen und Rule (2004) muss eine Auseinandersetzung stattfinden. Findet diese erziehungswissenschaftliche Auseinandersetzung allerdings kaum statt, da die Studierenden z.B. aufgrund des Curriculums nur wenig Berührungspunkte mit dem Fach haben (nur wenige Veranstaltungen besuchen), wäre gar nicht anzunehmen, dass die Überzeugungen sich über die Zeit weiterentwickeln.

Für diesen Erklärungsansatz sprechen auch die bestehenden Unterschiede in den epistemologischen Überzeugungen in Abhängigkeit von der Schulart. Trotz der unbalancierten Zellengröße und der geringen Anzahl an Förderschulstudierenden ($n$ = 32) zeigte sich ein signifikanter Unterschied (5,5% aufgeklärte Varianz) in der Überzeugung zur Absolutheit des Wissens zwischen diesen Förderschulstudierenden[2] und den Grundschulstudierenden ($n$ = 332). Die Förderschulstudierenden verfügen über sophistizierte Überzeugungen zur Absolutheit. Gleichzeitig handelt es sich bei den Förderschulstudierenden um die Gruppe von Studierenden, die curricular die meisten erziehungswissenschaftlichen Inhalte haben, d.h. die meiste Gelegenheit dazu haben, ihre Überzeugungen zum erziehungswissenschaftlichen Wissen weiterzuentwickeln. Die Gruppe der Gymnasiallehramtsstudierenden, unter denen sich die Mehrheit zusätzlich in höheren Fachsemestern befand, verfügte über sophistizierte epistemologische Überzeugungen im Vergleich zu den Grund-, Real-, Mittel- und

[2]Unterschiede zwischen den Förderschulstudierenden und den Grund-, Real-, Mittel- und Sekundarschulstudierenden lagen ebenfalls vor, wurden aber nicht signifikant.

Sekundarschulstudierenden. An der Uni Halle, an der viele ($n$ = 88 von 193) dieser Gymnasiallehramtsstudierenden studieren, sieht das Curriculum für Gymnasiallehramtsstudierende mehr erziehungswissenschaftliche Inhalte vor, als für die Studierenden der anderen Schularten (mit Ausnahme Förderschule). Hier liegen also Hinweise vor, dass eine Auseinandersetzung mit den erziehungswissenschaftlichen Inhalten zu sophistizierteren Überzeugungen führen kann. Allerdings zeigt sich auch, dass diese Auseinandersetzung insgesamt zu kurz kommt: Diese Ergebnisse, d.h. die universitätsübergreifend wenig fortgeschrittenen Überzeugungen zur Absolutheit, Einfachheit und Entwicklung des Wissens, und zwar sowohl für Studienanfänger als auch für fortgeschrittene Fachsemester, geben Hinweise darauf, dass neben den fachwissenschaftlichen Inhalten im Studium (dieser Universitäten) zu wenig erziehungswissenschaftliche Inhalte im Curriculum verankert sind und sich demnach zu wenig Lerngelegenheiten bieten, die es den Studierenden nicht nur ermöglichen würden, ihr Wissen in der Erziehungswissenschaft auszuweiten, sondern eben auch fortgeschrittene wissensbezogene Überzeugungen auszubilden.

Die Prüfung der Grundannahmen 1 bis 8 diente als Evidenzquelle um zu beurteilen, ob die Testwertinterpretationen als angemessen und gültig gelten können. Wie zuvor bereits beschrieben konnten die Grundannahmen 1 bis 6 angenommen werden, während Grundannahme 7 nicht angenommen und über Grundannahme 8 keine Aussage getroffen werden konnte. In Bezug auf den Validierungsprozess stellt sich die Tatsache, dass nur geringe Effektstärken für die Unterschiede in den Überzeugungen in Abhängigkeit des Standortes, der Schulart und des Studienfortschrittes gefunden werden konnten (Grundannahme 5 bis 8), zunächst als eine Einschränkung dar. Denn wie zuvor beschrieben, wurde aufgrund der theoretischen Überlegungen angenommen, dass diese Unterschiede bestehen müssten und sich demnach, wenn die Testwertinterpretationen gültig sind, mittels StEB abbilden lassen müssten. Insgesamt scheinen die epistemologischen Überzeugungen relativ wenig fortgeschritten zu sein, und Unterschiede zwischen den Standorten, Schularten und in Abhängigkeit des Studienfortschrittes zeigen sich kaum. Handelt es sich dabei allerdings um ein Resultat des geringen Umfangs der erziehungswissenschaftlichen Inhalte in den Studiengängen der betreffenden Universitäten (der womöglich Ausdruck der der Erziehungswissenschaft beigemessenen Bedeutung in der dortigen Lehrerbildung ist), kann das StEB Inventar natürlich keine Unterschiede finden, wo keine vorliegen. Wenn also der Erziehungswissenschaft curricular so wenig Platz eingeräumt wird und die Studierenden so wenig Gelegenheit erhalten, ihre erziehungswissenschaftlichen epistemologischen Überzeugungen weiterzuentwickeln, dann stellt es keine Einschränkung der Gültigkeit der Testwertinterpretationen des StEB dar, wenn der StEB kaum Unterschiede identifiziert. Zumal in den Ergebnissen der CFA über die signifikanten Varianzen der latenten Variablen des StEB gezeigt werden konnte, dass das StEB durchaus dazu in der Lage ist, bestehende Unterschiede in den epistemologischen Überzeugungen abzubilden. Auch die Tatsache, dass die Items trennscharf sind, zeigt, dass das StEB prinzipiell schon Unterschiede in den erziehungswissenschaftlichen epistemologischen Überzeugungen aufdecken könnte, wenn denn welche existierten.

An dieser Stelle muss betont werden, dass diese Ergebnisse nicht bedeuten, dass die

theoretisch abgeleiteten Grundannahmen 5 bis 8 in sich falsch sind. Die Grundannahmen 5 bis 8 sind unter der Voraussetzung richtig, dass Studierende in ihrem Studium erziehungswissenschaftliche Bildungserfahrungen machen, die ihnen die Gelegenheit zur Weiterentwicklung ihrer epistemologischen Überzeugungen geben. Und in so einem Fall, müsste die Ablehnung einer dieser Grundannahmen nach ihrer Prüfung möglicherweise zu der Schlussfolgerung führen, dass die zugehörigen Testwertinterpretationen ungültig wären. Im vorliegenden Fall trifft dies allerdings nicht zu. Wo keine Unterschiede vorliegen, kann von einem Erhebungsinstrument natürlich nicht erwartet werden, welche zu finden. Deshalb wäre es fehlerhaft, die Testwertinterpretationen auf Basis des Mangels an Unterschieden für ungültig zu erklären.

Dass die Beziehungen zwischen den StEB-Dimensionen und den Lernstrategiedimensionen des LIST-R nur rein korrelativ untersucht werden konnten, stellt eine weitere Einschränkung im Validierungsprozess des StEB dar. Die gefundenen Korrelationen sind alle nicht sehr hoch, aber trotzdem signifikant, wofür die Stichprobengröße ursächlich ist. Die Tatsache, dass diese Zusammenhänge signifikant wurden, ist demnach für die inhaltliche Interpretation nebensächlich. Bedeutsam ist vielmehr, dass sich diese gefundenen geringen Korrelationen alle konstruktkonform verhalten, wie in Abschnitt 8.3.4 (S. 289) hergeleitet wurde. Gleichzeitig zeigt sich auch, dass die Zusammenhänge zwischen den Lernstrategien und epistemologischen Überzeugungen in den bisherigen Studien nie sonderlich hoch ausfielen und die gefundenen Zusammenhänge daher mit den bisherigen Ergebnissen übereinstimmen. Zwischen den berichteten Studien, die die Beziehung zwischen Lernstrategien und epistemologischen Überzeugungen untersuchten, bestehen große Unterschiede bezüglich ihrer Qualität. Die bestehenden Studien, in denen die Zusammenhänge rein korrelativ untersucht wurden, wie die von Urhahne und Hopf (2004) und Urhahne und Harms (2006), sind insofern kritisch zu betrachten, als dass auch bei diesen davon auszugehen ist, dass die geringen Korrelationen allein aufgrund der Stichprobengröße signifikant wurden. Deshalb stellen die empirischen Studien, in denen die Beziehungen zwischen den Konstrukten mittels SEM untersucht wurden, eine geeignetere Basis dar.

Eine Möglichkeit wäre noch, den LIST-R nicht auf Itemebene in einem Strukturgleichungsmodell zu modellieren, sondern stattdessen Parcels, d.h. Gruppen von Items, aufzunehmen. Dies würde die Komplexität des Modells deutlich verringern und vermutlich zu einer besseren Modellpassung für den LIST-R führen.

In einer weiteren Studie sollten die Beziehungen zwischen den Lernstrategien und StEB-Dimensionen nochmals untersucht werden. Dabei wäre es angebacht, einen anderen Lernstrategiefragebogen einzusetzen, oder zumindest noch einen weiteren, bewährten Fragebogen neben dem LIST-R einzusetzen. Zusätzlich sollte gleichzeitig auch ein anderes Instrument zur Erfassung der epistemologischen Überzeugungen eingesetzt werden. So ließen sich die Ergebnisse des StEB Inventars mit den Ergebnissen des anderen Überzeugungsfragebogens vergleichen. Es könnte festgestellt werden, ob die Beziehungen, die die StEB-Skalen zu den Lernstrategieskalen aufweisen, mit den Beziehungen des anderen Überzeugungsfragebogens zu den Lernstrategieskalen vergleichbar sind. Da das StEB Inventar allerdings der erste Fragebogen ist, der speziell erziehungswissenschaftliche epistemologische

Überzeugungen erhebt, ist es schwierig, ein geeignetes anderes Erhebungsinstrument auszuwählen. Für den Vergleich würde sich meines Erachtens am ehesten der DEBQ von Hofer (2000) eignen, der auf der gleichen theoretischen Grundlage basiert.

Für die Überzeugungsskalen des finalen StEB Inventars wurden Reliabilitäten zwischen $\omega$ = .70 bis $\omega$ = .82 erreicht. Üblicherweise wird eine Reliabilität von über .80 als ausreichend hoch eingestuft (Bühner, 2011, S. 81). Wie Döring und Bortz (2016, S. 443) warnen, sollte diese Einstufung allerdings nicht unreflektiert angewendet werden. Vielmehr müssen bei der Interpretation der Reliabilität die Art des gemessenen Merkmals sowie mögliche methodische Alternativen berücksichtigt werden (Döring & Bortz, 2016, S. 443). Die existierenden Skalen zur Erfassung epistemologischer Überzeugungen verfügen üblicherweise über Reliabilitäten zwischen $\alpha$ .70 und .80 und liegen häufiger sogar unter .70 (siehe Kapitel 2.9.2, Tab. 2.13, 2.14, 2.15 und 2.16, ab S. 116). In Anbetracht dessen und in Anlehnung an Nunnally (1978) werden die erreichten Reliabilitätswerte des StEB als ausreichend hoch erachtet (vgl. auch Klopp, 2014; Pintrich, 2002).

## 9.5 Forschungspragmatische und praktische Implikationen des StEB Inventars

Der Beitrag dieser Arbeit besteht darin, das vorherrschende epistemologische Überzeugungssystem von Lehramtsstudierenden zum erziehungswissenschaftlichen Wissen über das StEB Inventar erfassbar zu machen. Mit dem StEB Inventar lassen sich die Überzeugungen von Lehramtsstudierenden zur Absolutheit, Einfachheit und Entwicklung des erziehungswissenschaftlichen Wissens sowie zur Multimodalität des erziehungswissenschaftlichen Wissenserwerbs aus Studientexten ermitteln.

Das StEB Inventar eignet sich dazu, interindividuelle Vergleiche anzustellen. Es kann folglich dazu genutzt werden, Aussagen darüber zu treffen, ob ein Lehramtsstudierender oder eine Gruppe Lehramtsstudierender im Vergleich zu einem anderen oder einer anderen Gruppe über weniger sophistizierte oder aber sophistiziertere epistemologische Überzeugungen zu erziehungswissenschaftlichem Wissen im Hinblick auf die verschiedenen Dimensionen verfügt. Über einen Pre-Post-Einsatz des StEB Inventars ist es außerdem möglich, intraindividuelle Veränderungen in den Ausprägungen epistemologischer Überzeugungen zu erziehungswissenschaftlichem Wissen aus Studientexten (zu den verschiedenen Dimensionen) zu identifizieren. Damit öffnet sich ein breites Feld möglicher Anwendungen in der Bildungsforschung, Lehrerbildung und Hochschulforschung. In der Hochschulforschung könnte das StEB Inventar beispielsweise eingesetzt werden, um zu prüfen, ob es mittels bestimmter Lehrkonzepte gelingt (z.B. solchen die auf die Reflexion der eigenen Überzeugungen zum erziehungswissenschaftlichen Wissen abzielen), eine epistemologische Entwicklung in Richtung eines sophistizierteren Überzeugungssystems zum erziehungswissenschaftlichen Wissen unter Lehramtsstudierenden zu fördern.

Auch zur zielgerichteten Förderung in der erziehungswissenschaftlichen Hochschullehre lässt sich das StEB Inventar einsetzen. Denn wenn die epistemologischen Überzeugungen zum erziehungswissenschaftlichen Wissen von einer Gruppe Studierender bekannt sind, dann können Hochschullehrende im Seminar zielgerichtet, je nach identifizierter zu fördernder Überzeugungsdimension beispielsweise:

- einen Fokus auf sich widersprechende erziehungswissenschaftliche Quellen legen.
- unterschiedliche Autoritäten über unterschiedliche erziehungswissenschaftliche Texte sprechen lassen.
- auf der Grundlage einer Annahme aus einer Tageszeitung die Originalforschung suchen und die allgemeinen und vereinfachten Annahmen der Journalisten demystifizieren.
- mit einem Forschungsnachweis zeigen, wie zu einfache Annahmen (z.B. zur Motivation ohne Kenntnis der Motivlagen oder zur Handlungsanregung ohne Kenntnis von Zielsetzungen) zu falschen und sogar schädlichen Entscheidungen führen.

- einfache Plausibilität kritisch beleuchten, indem anhand einer Reihe aufgeklärter Irrtümer in ein komplexes Modell geführt wird, um aufzuzeigen, dass schon die subjektive Theoriebildung besser wird, wenn sie immer komplexer wird.
- entlang von Videofallvignetten zeigen, wie sich Alltagsannahmen in unvorteilhaften Unterrichtssituationen abbilden.

Es wäre zu prüfen, ob das StEB Inventar in der Lehrerbildung als Werkzeug eingesetzt werden kann, um Lehramtsstudierenden dabei zu helfen, sich der eigenen Überzeugungen zum erziehungswissenschaftlichen Wissen bewusst zu werden und die bestehenden Überzeugungen zu reflektieren. Die Reflexion der eigenen wissensbezogenen Überzeugungen stellt einen wichtigen Bestandteil dabei dar, eine Weiterentwicklung des Überzeugungssystems zu fördern (Bendixen & Rule, 2004). Auch in Verbindung mit Praktika würde sich hier für die Lehramtsstudierenden die Möglichkeit bieten, die eigenen Überzeugungen hinsichtlich ihrer Auswirkungen auf das eigene Unterrichtshandeln zu reflektieren.

Im Hinblick auf die gefundenen Unterschiede in den Überzeugungen wäre ein systematischer Vergleich von Lehramtsstudierendengruppen von Interesse. Vor allem bei Lehramtsstudierenden für Förderschule und andere Lehramtsstudierendengruppen mit stärkerer curricularer Verankerung der erziehungswissenschaftlichen Inhalte im Studium wäre eine genauere Betrachtung der bestehenden epistemologischen Überzeugungen interessant. Hier wäre beispielsweise die längsschnittliche Entwicklung der erziehungswissenschaftlichen Überzeugungen über das Studium hinweg von Interesse. Außerdem lässt sich prüfen, inwieweit die erziehungswissenschaftlichen epistemologischen Überzeugungen im Studium (bei Studierendengruppen mit starker curricularer Verankerung der Erziehungswissenschaft) das pädagogische Handeln (z.B. Unterrichtsplanung, Unterrichtsstrategien) im Referendariat beeinflussen und das spätere unterrichtliche Handeln im Lehrberuf vorhersagen. Als eines der nächsten Vorhaben soll untersucht werden, inwieweit sich zum Ende des Studiums die epistemologischen Überzeugungen einer solchen Gruppe, die im Verlauf ihres Studiums zahlreichere erziehungswissenschaftliche Lerngelegenheiten hatte, von den epistemologischen Überzeugungen von Studierendengruppen unterscheiden, die weniger erziehungswissenschaftliche Lerngelegenheiten hatten und ob sich das pädagogische Handeln dieser Gruppen unterscheidet.

# A Eingesetzte StEB Testentwürfe

## StEB-Version der Expertenbefragung

Lehrstuhl für Erziehungswissenschaft mit
Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

## Instrumentenentwicklung: Epistemologische Überzeugungen zum Lernen mit Studientexten

**Hinweise zum Vorgehen für die Expertinnen und Experten:**

- Bitte lesen Sie sich zunächst die Beschreibungen der 5 Dimensionen (Subskalen) mit den jeweiligen Polen in Ruhe durch.
- Schätzen Sie danach bitte die Items in Bezug auf ihre Passung zu ihrer jeweiligen intendierten Dimension auf einer 11-stufigen Skala (von 1 = passt minimal zur Skala, 11 = passt maximal zur Skala) ein. Markieren Sie dazu die entsprechende Zahl jeweils mit der Texthervorhebungsfunktion gelb.
- Ab Seite 9 finden Sie eine Version des Fragebogens, wie er später bei den Studierenden eingesetzt werden soll. (Natürlich abzüglich der von Ihnen als ungeeignet eingeschätzten Items.) Sollten Ihnen noch Kritikpunkte in Bezug auf den Aufbau des Fragebogens, die Itemformulierung oder etwas anderes auffallen, freue ich mich sehr, wenn Sie mir diese als Kommentar anmerken.

**Hinweis zur Begriffsverwendung „Studientext":**

Unter Studientext werden in diesem Fragebogen alle wissenschaftlichen Texte verstanden, welche die Studierenden im Studienfach Pädagogik für Veranstaltungen und Prüfungen zum Lernen lesen. Das umfasst Texte, wie Forschungsartikel, Lehrbuchtexte, wissenschaftliche Beiträge, die von Hochschullehrenden in der Veranstaltung, im Seminar und auf Lernplattformen, wie ILIAS, Moodle und StudIP bereitgestellt werden.

## Beschreibungen der 5 Subskalen:

**Skala 1: Quelle des Wissens**

Die Dimension „Quelle des Wissens" betrifft die epistemologische Vorstellung über den Ursprung des Wissens. Das betrifft die Überzeugung, dass Wissen entweder eher von außen über Autoritäten an Lernende herangetragen wird oder eher auf der Fähigkeit der Lernenden aufbaut, Wissen in Interaktion mit anderen zu konstruieren (vgl. Hofer & Pintrich, 1997; Hofer, 2000)

**Mit den beiden Polen:**

| ***dualistisch/absolut (external)*** | ***multiplizistisch (internal)*** |
|---|---|
| Wissen wird vom Individuum als external (außerhalb des Selbst liegend) angesehen. Wissen liegt und hat seinen Ursprung in Autoritäten, die das Wissen an Personen übermitteln können.<br><br>**Bezug zu Studientexten**<br>Autoritäten sind in diesem Fall die Autorinnen und Autoren der Studientexte, die ihr Wissen in diesen Texten festhalten. Die Studientexte dienen dazu, das Wissen an die Studierenden zu übertragen, wenn diese die Studientexte lesen. | Das Individuum sieht in sich selbst den Ursprung von Wissen, in dem Sinne, dass in ihm selbst die Fähigkeit ruht Wissen in Interaktion mit anderen zu konstruieren.<br><br>**Bezug zu Studientexten**<br>Wissen kann einerseits selbst durch Auseinandersetzung mit dem Studientext und unter Berücksichtigung weiterer Texte konstruiert werden oder aber in Auseinandersetzung mit anderen, z.B. Kommilitonen. |

Lehrstuhl für Erziehungswissenschaft mit
Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

**Skala 2: Rechtfertigung des Wissens**

Die Dimension „Rechtfertigung des Wissens" umfasst, wie Lernende Behauptungen bewerten, von Belegen Gebrauch machen und wie sie Autorität und Expertise annehmen sowie Experteninnen und Experten bewerten. Individuen auf niedrigerem Level rechtfertigen Ansichten über Beobachtung, Autorität oder auf der Basis davon, was sich richtig anfühlt. Individuen auf höherer Stufe prüfen Behauptungen und wägen die Ansichten von Experten ab und integrieren sie (vgl. King & Kitchener, 1994; Hofer, 2000).

**Mit den beiden Polen**

| ***dualistisch/absolut*** | ***multiplizistisch/evaluativ*** |
|---|---|
| Rechtfertigung des Wissens erfolgt allein über Autorität oder Beobachtung. Richtig ist für das Individuum, was man beobachten kann, was Autoritäten annehmen oder – im Falle von Unsicherheit – was sich richtig anfühlt. Keine richtige Begründung für Meinungen und Positionen ist nötig und wird gegeben. | Wissen wird über Nachforschung und Prüfung begründet und gerechtfertigt. Das Individuum bewertet und integriert die Ansichten von Experten und neigt zu vielfältigen Meinungen und begründeten Urteilen. |
| **Bezug zu Studientexten**<br>Richtig ist, was man mal gelesen hat bzw. was die Autorinnen und Autoren berichten. Die Glaubwürdigkeit ist abhängig von Stellung, Position und Titel des Autors oder der Autorin. | **Bezug zu Studientexten**<br>Eine Meinung zum Studientext wird gebildet auf Basis kritischen Reflektierens der Textinhalte, genannter Forschungsergebnisse und im Abgleich mit dem Vorwissen und mit Positionen und Forschungsergebnissen anderer Autorinnen und Autoren. Es wird darauf Wert gelegt, dass die Autorinnen und Autoren logisch begründen, argumentieren und Gegenpositionen abwägen. |

**Skala 3: Sicherheit des Wissens**

Die Dimension „Sicherheit des Wissens" umfasst die beiden epistemologischen Vorstellungen, dass Wissen einerseits eher als feststehend angenommen wird, in dem Sinne, dass es „die eine richtige Antwort" (Wahrheit) gibt. Auf der anderen Seite steht die Überzeugung, dass es keine absoluten Wahrheiten gibt und Wissen mit Unsicherheit behaftet ist. Individuen auf diesem Level gehen davon aus, dass komplexe Probleme mehrere richtige Antworten haben können. Sie sind neuen Interpretationen gegenüber offen und auch der Möglichkeit, die eigene Ansicht im echten Austausch mit anderen zu verändern (Kuhn, 1991; King and Kitchener, 1994; Hofer, 2000, Conley et al., 2004).

**Mit den beiden Polen:**

| ***dualistisch/absolut*** | ***multiplizistisch/evaluativ*** |
|---|---|
| Wissen wird vom Individuum als feststehend wahrgenommen, d.h. es wird angenommen, dass es auf Fragen/Probleme ausschließlich eine richtige Antwort gibt. | Für das Individuum ist Wissen mit Unsicherheit behaftet und es gibt keine absoluten Wahrheiten. Es kann mehr als eine richtige Antwort auf Fragen oder Probleme geben. Neuen Interpretationen wird offen gegenüber gestanden, wie auch der Möglichkeit, die eigene Ansicht zu ändern. |
| **Bezug zu Studientexten**<br>Was im Studientext steht ist entweder wahr oder falsch. Es gibt nur eine richtige Antwort/Ansicht. | **Bezug zu Studientexten**<br>Mehrere im Studientext berichtete Positionen könnten zutreffen oder nicht zutreffen. Was richtig ist, kann auch vom Kontext abhängig sein. Manche Ansichten treffen vielleicht wahrscheinlicher zu als andere. Es gibt keine absolute Wahrheit, deshalb ist der Austausch von Ansichten wichtig. |

Lehrstuhl für Erziehungswissenschaft mit
Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

**Skala 4: Entwicklung des Wissens**

Die Dimension „Entwicklung des Wissens" umfasst die epistemologischen Überzeugungen zur Veränderlichkeit des Fachbereichs (d.h. z.B. Pädagogik als Disziplin). Auf der einen Seite steht die Auffassung, dass die Fachwissenschaft als solche nicht veränderlich ist, d.h. dass sich nicht ändert, was jetzt als wahr angenommen wird. Auf der anderen Seite wird die Disziplin als sich entwickelnd und veränderlich wahrgenommen. Auf der Basis neuer Entdeckungen, Studien und Forschungsergebnisse ändern Fachexpertinnen und -experten ihre Ansichten darüber, was in ihrem Fach zutrifft (vgl. Conley et al., 2004).

**Mit den beiden Polen:**

| ***dualistisch/absolut (isolated)*** | ***multiplizistisch (integrated)*** |
|---|---|
| Das Individuum ist überzeugt, dass die Disziplin sich nicht über die Zeit verändert, sondern statisch ist. Was Fachexpertinnen und -experten heute als wahr annehmen, wird auch in Zukunft als wahr angenommen werden. | Die Disziplin wird als über die Zeit veränderlich wahrgenommen. Neue Entdeckungen, neue Forschung und Forschungsergebnisse führen dazu, dass Fachexpertinnen und -experten ihre Ansichten darüber ändern, was zutrifft. |
| **Bezug zu Studientexten**<br>Die in Studientexten vertretenen Ansichten verändern sich nicht über die Zeit, d.h. neuere und ältere Texte vertreten dieselben Positionen. Autorinnen und Autoren bleiben bei ihren Positionen. | **Bezug zu Studientexten**<br>Die in den Studientexten vertretenen Ansichten über das Fach und darüber, was im Fach zutrifft, können sich über die Zeit aufgrund von Forschung verändern. Die Ansichten und Positionen von Autorinnen und Autoren können sich ändern. |

**Skala 5: Struktur des Wissens**

Die Dimension „Struktur des Wissens" betrifft die epistemologische Vorstellung über die Struktur von Wissen. Wissen wird auf einem Kontinuum betrachtet, das von Wissen als Anhäufung von Fakten bis hin zu Wissen als in gegenseitiger Beziehung stehende Konzepte reicht. Auf niedrigerem Level wird Wissen als diskrete (klar abgrenzbare), konkrete zu wissende Fakten betrachtet, d.h. Wissen ist die Anhäufung von Fakten. Auf höherer Ebene wird Wissen als hoch vernetzte, komplexe Konzepte betrachtet (Schommer, 1990, 1994).

**Mit den beiden Polen:**

| ***dualistisch/absolut (isolated)*** | ***multiplizistisch (integrated)*** |
|---|---|
| Wissen wird vom Individuum als diskrete (klar abgrenzbare), konkrete zu wissende Fakten betrachtet. Wissen stellt eine Anhäufung von Fakten dar. | Wissen wird als hoch vernetzte, komplexe Konzepte wahrgenommen. Wissen ist demnach nicht einfach sondern komplex. |
| **Bezug zu Studientexten**<br>Die Fakten in einem Studientext stellen das Wissen dar, das dieser vermitteln kann. | **Bezug zu Studientexten**<br>Die Fakten, die ein Studientext enthält, sind nicht auch gleichzeitig das Wissen, das dieser enthält. Da Wissen komplex und hoch vernetzt ist, können Fakten nur unvollständiges Wissen sein. Verschiedene Positionen zu einem Thema müssen berücksichtigt werden, um etwas wissen zu können. |

Lehrstuhl für Erziehungswissenschaft mit
Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

| **1. Skala: Quelle des Wissens** | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Einschätzung über Passung zur Subskala** | **Minimal** | | | | | | | | | | **Maximal** |
| qw_1 | Was in einem Studientext steht, kann man auch genau so einsetzen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_2 | Expertinnen und Experten schreiben nur vollständig gesicherte Erkenntnisse in Studientexte. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_3 | Bevor etwas in einem Studientext landet, wird sichergestellt, dass das Wissen auch wahr ist. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_4 | Expertinnen und Experten unterliegen strengen Regeln, die für uns sicherstellen, dass das Wissen in Studientexten weitgehend wahr ist. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_5 | Auch die besten Expertinnen und Experten irren sich häufiger und müssen ihre Ansichten später korrigieren. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_6 | Es ist fast egal, wie angesehen eine Autorin oder ein Autor ist. Es kann trotzdem Falsches in den Studientexten stehen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_7 | Lesende von Studientexten müssen über eigene Methoden der Überprüfung verfügen, um die Reichweite des Wissens darin einschätzen zu können. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_8 | Für die Anwendungsbereiche im Studium reicht es aus, wenn man Inhalte von Studientexten im Großen und Ganzen einfach mal glaubt. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_9 | Angesehene Studientextautorinnen und -autoren gewinnen ihr Ansehen dadurch, dass sie viel öfter richtig liegen als andere. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_10 | Es ist besser, sich bei Studientexten an die großen Autoritäten im Fach zu halten. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_11 | Für mein eigenes Wissen ist es wichtig, mir über möglichst vielen verschiedenen Ansichten einen Überblick zu verschaffen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_12 | Meine eigenen Ansichten erlange ich aus sehr vielen verschiedenen Quellen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_13 | Ich bin darauf angewiesen, dass mir jemand sagt, was genau gilt, damit ich mich dann später daran halten kann. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_14 | Wenn es in einem Bereich keine anerkannten Expertinnen und Experten gibt, dann muss ich mich eher auf meine Intuition und Erfahrung verlassen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_15 | Meine eigene Erfahrung vermag mehr Wahrheit zu erkennen als die meisten theoretischen Texte. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_16 | Für die Vertrauenswürdigkeit des enthaltenen Wissens spielt es eine wichtige Rolle, wann ein Studientext publiziert wurde. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_17 | Gute Studientexte sind diejenigen, mit denen man die Prüfungen einfach und gut besteht. Sie enthalten das richtige Wissen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_18 | Wenn ich in einer Prüfung eine sehr gute Note bekomme, dann habe ich auch genau die wichtigen Sachen aus den Studientexten gelernt. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_19 | Wenn eine Professorin oder ein Professor einer Studientextmeinung zustimmt, dann ist sie wahrscheinlich richtig. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_20 | Was in Studientexten steht, muss man glauben. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_21 | Was man in einem Studientext liest, ist sicher wahr. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_22 | Nur die Autorinnen und Autoren der Studientexte wissen genau, was in ihrem Fach zutrifft. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_23 | Im Studium muss man sich darauf verlassen, dass das, was man in Studientexten liest, zutrifft. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_24 | Neues Wissen erlangt man, indem man es sich u.a. unter Zuhilfenahme von Studientexten konstruiert. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_25 | Studierende sollten nicht anzweifeln, was in Studientexten steht. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_26 | Studierende sollten die Autorität der Autorin oder des Autors eines Studientextes nicht in Frage stellen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_27 | Wenn ein Studientext ein bestimmtes Vorgehen empfiehlt, halte ich mich daran. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_28 | Man kann fast alles glauben, was man in Studientexten liest. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_29 | Wenn mir beim Lesen eines Studientextes auffällt, dass dieser meiner Erfahrung nicht entspricht, dann muss ich mich wohl geirrt haben. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_30 | Ich bin mir sehr sicher etwas zu wissen, wenn ich die in einem Studientext beschriebene Position kenne. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| qw_31 | Manchmal muss man die in einem Studientext gegebene Antwort einfach akzeptieren, auch wenn man sie nicht versteht. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |

Lehrstuhl für Erziehungswissenschaft mit
Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

**2. Skala: Rechtfertigung des Wissens**

| | Einschätzung über Passung zur Subskala | Minimal | | | | | | | | | | Maximal |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| rw_1 | In Studientexten dargestelltes Wissen muss mir in erster Linie plausibel sein. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_2 | Mir ist es sehr wichtig, dass sich die Inhalte von Studientexten für mich richtig anfühlen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_3 | Es ist weniger wichtig, ob mir etwas in Studientexten sofort plausibel erscheint. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_4 | Es ist mir wichtig, dass die Studientextmeinung durch ausgewiesene Autoritäten legitimiert ist. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_5 | Ich muss nachvollziehen können, wie genau eine Erkenntnis im Studientext zustande kam, um ihr Glauben zu schenken. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_6 | Mit den richtigen Methoden kann auch jemand, der keine große Autorität ist, zu bahnbrechenden Erkenntnissen gelangen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_7 | Für mich ist es schon wichtig, wer etwas geschrieben hat, wenn ich dem Inhalt Glauben schenken soll. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_8 | Studientexte werden von meinen Hochschullehrenden korrekt ausgewählt, so dass ich mich auf sie verlassen kann. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_9 | Es ist Aufgabe meiner Hochschullehrenden, die richtigen Studientexte für mich auszuwählen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_10 | Ich suche oft nach weiteren Quellen, weil mir die von meinen Hochschullehrenden ausgewählten Studientexte nicht ganz ausreichen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_11 | Ich ziehe neben Studientexten noch weiteren Quellen zurate, um verschiedene Forschungsmeinungen zu konsultieren. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_12 | Es ist als Studierende/r eher meine Aufgabe, die für mein Studium geeigneten Studientextquellen auszuwählen und zu bewerten. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_13 | Mir selbst kommt bei der Gestaltung meines Wissens eine wesentliche Aufgabe zu, weil ich mich nicht einfach auf einzelne Studientexte verlassen kann. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_14 | Die Auswahl und Gestaltung meines zu erwerbenden Wissens ist Aufgabe und Pflicht der Hochschullehrenden. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_15 | Für meinen Wissenserwerb ist es notwendig, dass meine Hochschullehrenden durch Qualifikation ausgewiesene Expertinnen und Experten auf ihrem Fachgebiet sind. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_16 | Hausarbeiten meiner Mitstudierenden sind wegen der geringeren Qualifikation in der Regel schlechter als die von Professorinnen und Professoren verfassten Lehrwerke. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_17 | Ein Studientext mit einem guten Inhalt wird üblicherweise auch von einem guten Verlag publiziert. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_18 | Ein Studientext aus einem auflagenstarken Lehrbuch spricht für einen vertrauenswürdigen Inhalt. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_19 | Wichtige Inhalte von Studientexten sind in der Regel auch spontan plausibel. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_20 | Mit genügend Zeit und Ressourcen könnte ich die meisten Forschungen und Studien aus meinen Studientexten auch selbst durchführen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_21 | Gute Ideen können von jedem kommen, nicht nur von Autorinnen und Autoren der Studientexte. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_22 | Gute Theorien in Studientexten erkennt man daran, dass sie sich auf viele verschiedene Forschungsergebnisse stützen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_23 | Um einschätzen zu können, ob die in einem Studientext dargestellte Meinung verlässlich ist, versuche ich Verbindungen zu meinem Vorwissen herzustellen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_24 | Wenn ich einen Studientext lese, versuche ich, mir ein eigenes Erklärungsmodell zu bauen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_25 | Um beim Lesen eines Studientextes zu echten Erkenntnissen zu gelangen, versuche ich einzuschätzen, ob das was ich lese, in sich logisch ist. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_26 | Den in einem Studientext aufgestellten Behauptungen kann man erst trauen, wenn man sich mit mehreren Wissensquellen zu dem Thema vertraut gemacht hat. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_27 | Ich verstehe einen Studientext besser, wenn ich ihn nicht nur lese, sondern selbst durchdenke. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_28 | Ich bilde mir meine Meinung zu Themen nicht auf Basis eines einzelnen Studientextes. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |

Lehrstuhl für Erziehungswissenschaft mit
Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| rw_29 | Wenn die Autorin oder der Autor eines Studientextes eine hohe Position im Fachgebiet inne hat, dann wird schon zutreffen, was sie/er schreibt. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_30 | Ich vertraue den Inhalten von Studientexten, auch wenn die Autorinnen und Autoren nicht Begründungen und Belege anführen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_31 | Wenn eine Autorin oder ein Autor nur Behauptungen aufstellt, ohne diese zu belegen oder gut zu begründen, dann halte ich den Studientext für weniger vertrauenswürdig. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_32 | Nicht Studientexte, sondern eigene Erfahrung ist der beste Weg, um etwas in diesem Fach herauszufinden. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_33 | Da ich über keine Möglichkeit verfüge zu prüfen, ob stimmt was ich in den Studientexten lese, kann ich es genauso gut einfach glauben. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| rw_34 | Ich akzeptiere eher die Ansichten von Praktikerinnen und Praktikern des Fachs als die Ansichten von den Forschenden, die die Studientexte verfassen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |

| **3. Skala: Sicherheit des Wissens** | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Einschätzung über Passung zur Subskala** | **Minimal** | | | | | | | | | | **Maximal** |
| gw_1 | Die Studientexte enthalten stets gesichertes Wissen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_2 | Forschende sollten sich erstmal einig über ihre Aussagen werden, bevor sie diese in Studientexten anführen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_3 | Es ist Aufgabe der Expertinnen und Experten, die sachlichen Differenzen im Wissen auszudiskutieren. Erst danach sollten sie Studientexte verfassen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_4 | Sind Forschungserkenntnisse erst einmal durch Studien bewiesen, kann man sich auch darauf verlassen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_5 | In meinem Studium bin ich darauf angewiesen, dass mir gesicherte Erkenntnisse in Studientexten vorgestellt werden. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_6 | Ich kann nicht alles überprüfen, deswegen muss ich eben glauben, was in den Studientexten steht. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_7 | Ich erwarte von meinen Studientexten gesichertes Wissen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_8 | Studientexte sollten immer auch ungesichertes Wissen enthalten. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_9 | Studientexte sollten neben Herleitungen und Aussagen auch Auskunft darüber geben, wie die einzelnen Aspekte genau erforscht wurden. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_10 | Mir ist nicht so wichtig, wie man genau zu den Inhalten in den Studientexten kam. Wichtig ist eher, was dabei heraus kommt und ob ich das anwenden kann. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_11 | Beim Lesen eines Studientextes ist es das Wichtigste zu erkennen, welche der dargestellten Lösungen die richtige ist. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_12 | Wenn in einem Studientext zwei verschiedene Meinungen zu einem Thema dargestellt werden, kann maximal nur eine davon richtig sein. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_13 | Ein Studientext sollte die richtige Antwort enthalten, statt viele Ansichten zu diskutieren. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_14 | Die meisten Autorinnen und Autoren, stellen die Themen komplizierter dar, als sie eigentlich sind. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_15 | Was in einem Studientext steht, trifft in jedem Fall zu. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_16 | Es kommt bei einem Studientext darauf an, dass viele unterschiedliche Ansichten diskutiert werden. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_17 | Wenn in einem Studientext eine Ansicht vertreten wird, die ich bisher noch nicht kannte, dann bin dafür offen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| gw_18 | Wenn in einem Studientext mehrere unterschiedliche Ansichten beschrieben werden, dann kann je nach Kontext variieren, welche Sichtweise gerade die bessere ist. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |

Lehrstuhl für Erziehungswissenschaft mit
Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

| **4. Skala: Entwicklung des Wissens** | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Einschätzung über Passung zur Subskala** | **Minimal** | | | | | | | | | | **Maximal** |
| ew_1 | Die wichtigsten Anwendungsaspekte des Wissens ändern sich eigentlich nicht sonderlich über die Zeit. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_2 | Es ist völlig normal, dass Expertinnen und Experten sich über viele Bereiche uneins sind: Wissen wird sich immer stark verändern. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_3 | Wie etwas verstanden wird, kann sich in der Forschung sehr schnell ändern. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_4 | Es ist wahrscheinlich, dass der Großteil der Erkenntnisse in den Studientexten in zehn Jahren ganz anders aufgefasst wird. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_5 | Bei Studientexten kommen über die Jahre zwar ein paar Dinge hinzu aber dennoch bleiben die Erkenntnisse weitestgehend erhalten. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_6 | Es reicht aus, dass ich mich jetzt mit den Studientexten auseinandersetze. Abgesehen von Kleinigkeiten, werde ich das meiste daraus auch noch in zehn Jahren einsetzen können. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_7 | In unseren Fachbereichen reicht es aus, die Inhalte einmal gut verstanden zu haben. Es ist nicht notwendig, sich alle paar Jahre nochmal komplett neu darin zu orientieren. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_8 | Ich werde mich schon in wenigen Jahren den meisten Inhalten aus den Studientexten neu widmen müssen, weil viele Dinge wieder überholt sein werden. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_9 | Heutzutage finden Forschende eher nur sehr spezielle Kleinigkeiten an neuem Wissen heraus. Im Großen und Ganzen bleiben die Inhalte aber über die Jahrzehnte in etwa gleich. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_10 | Forschende entdecken sehr oft Relevantes und Neues, so dass ich mein Wissen in kürzerer Zeit immer wieder aktualisieren muss. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_11 | Wenn sich Erkenntnisse in einem Inhaltsbereich oft verändern, sollten sie nicht in Studientexten vorkommen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_12 | Einige Ansichten in Studientexten entsprechen nicht mehr dem, was früher als richtig angenommen wurde. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_13 | Die Vorstellungen, die in Studientexten dargestellt werden, verändern sich über die Zeit. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_14 | Es gibt manche Fragen, die auch in Studientexten nicht beantwortet werden. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_15 | Manchmal ändern Autorinnen und Autoren ihre Meinung darüber, was in ihrem Fach wahr ist. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_16 | Was heute in Studientexten als wahr gilt, trifft in Zukunft auch noch zu. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| ew_17 | Durch neue Entdeckungen kann sich verändern, was die Autorinnen und Autoren von Studientexten für wahr halten. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |

Lehrstuhl für Erziehungswissenschaft mit
Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

| | **5. Skala: Struktur des Wissens** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Einschätzung über Passung zur Subskala** | **Minimal** | | | | | | | | | | **Maximal** |
| sw_1 | Die Dinge sind einfacher als sie in den meisten Studientexten dargestellt werden. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_2 | Wenn es nicht gelingt, etwas in Studientexten einfach dazustellen, dann ist es wahrscheinlich falsch. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_3 | Wenn Autorinnen und Autoren etwas komplex darstellen, dann versuchen sie damit oft zu verbergen, dass sie sich eigentlich gar nicht so gut auskennen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_4 | Wenn etwas komplex ist, dann führt auch kein Weg daran vorbei, diesen Sachverhalt auch in seiner ganzen Komplexität im Studientext darzustellen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_5 | Studientexte dürfen auf keinen Fall widersprüchlich sein. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_6 | Bei gehäuften Widersprüchen in Studientexten ist eigentlich immer etwas faul. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_7 | Bei guten Studientexten wird streng nach einzelnen Inhalten getrennt. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_8 | Bei guten Studientexten wird vermieden, dass sich die verschiedenen Bereiche zu sehr überlappen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_9 | Gute Studientexte verbinden viele Inhaltsbereiche und führen diese zusammen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_10 | Gute Studientexte gehen auf die verschiedenartigsten Verbindungen zu inhaltlichen Nachbarfeldern ein. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_11 | Nützliches Wissen lässt sich in Studientexten immer auch sehr einfach darstellen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_12 | Sobald Wissen aus Studientexten nützlich wird, ist es eigentlich auch immer recht komplex. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_13 | Nützliches Studientextwissen ist wie ein Rezept organisiert: Es werden in der Reihenfolge der Handlungen alle wesentlichen Schritte genannt. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_14 | Weil die echte Welt so vielfältig ist, ist auch das Wissen darüber komplex und sogar oft widersprüchlich. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_15 | In Studientexten werden Dinge oft unnötig verkompliziert. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_16 | Die echte Welt ist viel komplexer als das in unseren Studientexten oft dargestellt wird. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_17 | Wissen aus Studientexten muss immer einfach anzuwenden sein. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_18 | Wirklich wirksames Wissen muss eher komplex dargestellt werden. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_19 | Echtes Wissen ist auch in Studientexten immer widerspruchsfrei. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_20 | Brauchbares Wissen wird in Studientexten immer in einfache Ursache-Wirkung-Beziehungen herunter gebrochen. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_21 | Wenn man sich die Fakten aus den Studientexten aneignet, weiß man, was es zu wissen gibt. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_22 | Ich versuche die Informationen eines Studientextes mit denen aus anderen Studientexten zu verbinden. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_23 | Beim Lesen eines Studientextes suche ich nach den spezifischen Fakten. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_24 | Einen Studientext verstehen heißt, eher das große Ganze zu verstehen, als die Details. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_25 | Ein Studientext ist gut, wenn er die wichtigsten Fakten zum Thema enthält. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| sw_26 | Ein Studientext sollte alle Fakten enthalten, die zu dem Themenbereich wichtig sind. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |

**Literatur**

Conley, A. M., Pintrich, P. R., Vekiri, I. & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology* 29, 186-204.

Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology*, 25(4), 378–405.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88–140.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology,* 82, 498–504.

Schommer, M. (1994). An emerging conceptualization of epistemological beliefs and their role in learning. In R. Garner & P. A. Alexander (Eds.), *Beliefs about text and instruction with text.* Hillsdale, NJ: Erlbaum.

## StEB-Version der Studie 1

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

# Einstellungen zu pädagogischen Studientexten

In diesem Fragebogen geht es um Ihre Ansichten zu Studientexten im Studienbereich Pädagogik, Schulpädagogik und Erziehungswissenschaft.

## Allgemeine Angaben

Bitte tragen Sie die ersten drei Buchstaben des Vornamens Ihrer Mutter, die ersten drei Buchstaben des Vornamens Ihres Vaters, Ihren Geburts**tag** sowie die ersten drei Buchstaben Ihres Geburtsortes als Code ein:

| Persönlicher Code | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (Vorname Mutter, z.B. KER für Kerstin) | | | (Vorname Vater, z.B. MIC für Michael) | | | (Geburts**tag**, z.B. 06 für 6. Mai 1995) | | | (Geburtsort, z.B. PAS für Passau) | | |

| Alter | | Geschlecht | W ☐ |
|---|---|---|---|
| | | | M ☐ |

| Name Ihrer Hochschule | |
|---|---|

| Lehramt<br>Bachelor<br>Master | ☐<br>☐ | Studierte Schulart | (z.B. Grundschule) |
|---|---|---|---|

| Studiengang | |
|---|---|

| Fächerkombination im Lehramt | (z.B. Englisch und Deutsch) |
|---|---|

| Anzahl der Fachsemester | | *(Anzahl der Semester in Ihrem aktuellen Studienfach.)* |
|---|---|---|
| Anzahl der Hochschulsemester | | *(Anzahl der Semester, die Sie bereits insgesamt an einer Hochschule eingeschrieben sind.)* |

| Studienbeginn | WiSe ____/____ *(bitte Jahre eintragen)*<br>oder<br>SoSe _____ *(bitte Jahr eintragen)* |
|---|---|

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

**Wichtige Hinweise zum Ausfüllen des folgenden Fragebogens:**

Bitte lesen Sie sich alle Aussagen sorgfältig durch und entscheiden Sie jeweils auf einer Skala von 1 (trifft gar nicht zu) bis 5 (trifft voll zu), inwieweit Sie der Aussage zustimmen bzw. inwieweit die Aussage auf Sie zutrifft.

Bitte beziehen Sie die Aussagen auf die Studienbereiche Pädagogik, Schulpädagogik und Erziehungswissenschaft.

Bitte geben Sie ehrlich Ihre Meinung an. Es gibt keine falschen Antworten. Richtig ist, was Sie denken und Ihre Daten werden vollständig vertraulich behandelt.

Es ist außerdem wichtig, dass Sie alle Aussagen auf dem Bogen beantworten. Bitte lassen Sie also keine Aussage weg, auch wenn Sie sich nicht sofort für eine Antwort entscheiden können.

Bitte machen Sie nur ein Kreuz je Aussage.

**Wichtig!**

Unter **pädagogischen Studientexten** werden in diesem Fragebogen alle wissenschaftlichen Texte verstanden, die Sie in den Studienbereichen **Erziehungswissenschaft, Pädagogik und Schulpädagogik** für Ihr Studium lesen.

Das umfasst sowohl Texte, die Sie sich selbst suchen sowie solche, die Ihnen von Hochschullehrenden für Veranstaltungen (z.B. Vorlesung, Seminar, Übung) bereitgestellt werden (z.B. auf ILIAS, Moodle oder StudIP).

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

## *Lernen mit Studientexten*

| | Wie stark stimmen Sie den Aussagen zu in Bezug auf Ihren **Studienbereich Pädagogik/Schulpädagogik/Erziehungs-wissenschaft**? | **trifft gar nicht zu** | **trifft eher nicht zu** | **teils teils** | **trifft eher zu** | **trifft voll zu** |
|---|---|---|---|---|---|---|
| 1 | Wenn mir beim Lesen eines pädagogischen Studientextes auffällt, dass dieser meiner Erfahrung nicht entspricht, dann bin ich mir sicher, mich wohl geirrt zu haben. | ① | ② | ③ | ④ | ⑤ |
| 2 | Ich bin der Überzeugung, dass Wissen in pädagogischen Studientexten nützlich ist, wenn es wie ein Rezept organisiert ist: Es werden in der Reihenfolge der Handlungen alle wesentlichen Schritte genannt. | ① | ② | ③ | ④ | ⑤ |
| 3 | Es ist nicht so wichtig, wie man genau zu den Inhalten in den pädagogischen Studientexten kam. Wichtig ist eher, was dabei heraus kommt und ob ich das anwenden kann. | ① | ② | ③ | ④ | ⑤ |
| 4 | Auch die besten pädagogischen Studientextautorinnen und -autoren irren sich häufiger. | ① | ② | ③ | ④ | ⑤ |
| 5 | Wenn ein Sachverhalt komplex ist, dann führt auch kein Weg daran vorbei, ihn auch in seiner ganzen Komplexität im pädagogischen Studientext darzustellen. | ① | ② | ③ | ④ | ⑤ |
| 6 | Für mich ist es schon sehr wichtig, welche Person einen pädagogischen Studientext geschrieben hat, wenn ich dem Inhalt Glauben schenken soll. | ① | ② | ③ | ④ | ⑤ |
| 7 | Forschende sollten sich erstmal einig über ihre Aussagen werden, bevor sie diese in pädagogischen Studientexten anführen. | ① | ② | ③ | ④ | ⑤ |
| 8 | Wenn die Autorin oder der Autor eines pädagogischen Studientextes eine hohe Position im Fachgebiet inne hat, dann bin ich mir sicher, dass zutrifft, was sie/er schreibt. | ① | ② | ③ | ④ | ⑤ |
| 9 | Neues Wissen erlangt man in der Pädagogik, indem man es sich selbst u.a. unter Zuhilfenahme von Studientexten konstruiert. | ① | ② | ③ | ④ | ⑤ |
| 10 | Es ist wichtig, die Informationen eines pädagogischen Studientextes mit denen aus anderen pädagogischen Studientexten zu verbinden. | ① | ② | ③ | ④ | ⑤ |
| 11 | Ich bilde mir meine Meinung zu pädagogischen Themen nicht auf Basis eines einzelnen Studientextes. | ① | ② | ③ | ④ | ⑤ |
| 12 | Bevor etwas in einem pädagogischen Studientext landet, wird sichergestellt, dass das Wissen auch wahr ist. | ① | ② | ③ | ④ | ⑤ |
| 13 | Die Dinge sind einfacher als sie in den meisten pädagogischen Studientexten dargestellt werden. | ① | ② | ③ | ④ | ⑤ |
| 14 | Gute pädagogische Studientexte sind diejenigen, mit denen man die Prüfungen einfach und gut besteht. | ① | ② | ③ | ④ | ⑤ |
| 15 | Wenn in einem pädagogischen Studientext mehrere unterschiedliche Ansichten beschrieben werden, dann kann je nach Kontext variieren, welche Sichtweise gerade die bessere ist. | ① | ② | ③ | ④ | ⑤ |
| 16 | Ich stelle die Autorität der Autorin oder des Autors eines pädagogischen Studientextes nicht in Frage. | ① | ② | ③ | ④ | ⑤ |

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

| Wie stark stimmen Sie den Aussagen zu in Bezug auf Ihren **Studienbereich Pädagogik/Schulpädagogik/Erziehungswissenschaft**? | | **trifft gar nicht zu** | **trifft eher nicht zu** | **teils teils** | **trifft eher zu** | **trifft voll zu** |
|---|---|---|---|---|---|---|
| 17 | Da ich nicht prüfen kann, ob stimmt was ich in pädagogischen Studientexten lese, glaube ich es erstmal. | ① | ② | ③ | ④ | ⑤ |
| 18 | Wenn man sich die Fakten aus einem pädagogischen Studientext aneignet, dann weiß man das Wichtigste. | ① | ② | ③ | ④ | ⑤ |
| 19 | Ich verlasse mich darauf, dass in einem pädagogischen Studientext steht, was in meinem Fach genau gilt und daran halte ich mich. | ① | ② | ③ | ④ | ⑤ |
| 20 | Es ist mir wichtig einen pädagogischen Studientext nicht nur zu lesen, sondern selbst zu durchdenken. | ① | ② | ③ | ④ | ⑤ |
| 21 | Was man in einem pädagogischen Studientext liest, ist in der Regel wahr. | ① | ② | ③ | ④ | ⑤ |
| 22 | Das Wissen, das in pädagogischen Studientexten behandelt wird, ist komplex und sogar teilweise widersprüchlich, weil die echte Welt so vielfältig ist. | ① | ② | ③ | ④ | ⑤ |
| 23 | Ich verlasse mich eher auf meine eigenen Erfahrungen als auf Expertenmeinungen aus pädagogischen Studientexten. | ① | ② | ③ | ④ | ⑤ |
| 24 | Es ist als Studierende/r wichtig, die für mein Studium geeigneten pädagogischen Studientextquellen auszuwählen und zu bewerten. | ① | ② | ③ | ④ | ⑤ |
| 25 | Ich bin der Überzeugung, dass sich wirklich nützliches Wissen in pädagogischen Studientexten in der Regel sehr einfach darstellen lässt. | ① | ② | ③ | ④ | ⑤ |
| 26 | Ich verlasse mich darauf, dass mir die wichtigen Inhalte von pädagogischen Studientexten in der Regel spontan plausibel sind. | ① | ② | ③ | ④ | ⑤ |
| 27 | Für mein eigenes Wissen ist es wichtig, mir über möglichst viele verschiedene Ansichten aus verschiedenen pädagogischen Studientexten einen Überblick zu verschaffen. | ① | ② | ③ | ④ | ⑤ |
| 28 | Es ist wichtig, dass pädagogische Studientexte viele Inhaltsbereiche verbinden und diese zusammenführen. | ① | ② | ③ | ④ | ⑤ |
| 29 | Es ist wichtig, dass pädagogische Studientexte neben Herleitungen und Aussagen auch Auskunft darüber geben, wie die einzelnen Aspekte genau erforscht wurden. | ① | ② | ③ | ④ | ⑤ |
| 30 | Es ist wahrscheinlich, dass der Großteil der Erkenntnisse in den pädagogischen Studientexten in zehn Jahren ganz anders aufgefasst wird. | ① | ② | ③ | ④ | ⑤ |
| 31 | Ich verlasse mich auf Forschungserkenntnisse in pädagogischen Studientexten, wenn zugehörige Studien genannt werden. | ① | ② | ③ | ④ | ⑤ |
| 32 | Ich glaube generell schon, was in pädagogischen Studientexten steht. | ① | ② | ③ | ④ | ⑤ |
| 33 | Wie etwas verstanden wird, kann sich in der pädagogischen Forschung sehr schnell ändern. Demnach ändert sich dann auch, was in Studientexten berichtet wird. | ① | ② | ③ | ④ | ⑤ |

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

| | Wie stark stimmen Sie den Aussagen zu in Bezug auf Ihren **Studienbereich Pädagogik/Schulpädagogik/Erziehungs-wissenschaft**? | **trifft gar nicht zu** | **trifft eher nicht zu** | **teils teils** | **trifft eher zu** | **trifft voll zu** |
|---|---|---|---|---|---|---|
| 34 | Wenn ich einen pädagogischen Studientext lese, ist es mir wichtig, mir ein eigenes Erklärungsmodell der Inhalte zu bauen. | ① | ② | ③ | ④ | ⑤ |
| 35 | Es ist wichtig, dass pädagogische Studientexte auf die verschiedenartigen Verbindungen zu inhaltlichen Nachbarfeldern eingehen. | ① | ② | ③ | ④ | ⑤ |
| 36 | Ich zweifle nicht an, was in pädagogischen Studientexten steht. | ① | ② | ③ | ④ | ⑤ |
| 37 | Es ist wichtig, dass pädagogische Studientexte die verschiedenen Sachverhalte streng nach einzelnen Inhalten getrennt darstellen. | ① | ② | ③ | ④ | ⑤ |
| 38 | Den in einem pädagogischen Studientext aufgestellten Behauptungen kann man erst trauen, wenn man sich mit mehreren Wissensquellen zu dem Thema vertraut gemacht hat. | ① | ② | ③ | ④ | ⑤ |
| 39 | Nur die Autorinnen und Autoren der pädagogischen Studientexte wissen genau, was in ihrem Fach zutrifft. | ① | ② | ③ | ④ | ⑤ |
| 40 | Ein pädagogischer Studientext sollte die richtige Antwort, d.h. Fakten enthalten, statt viele Ansichten und Theorien zu diskutieren. | ① | ② | ③ | ④ | ⑤ |
| 41 | Die Vorstellungen, die in pädagogischen Studientexten dargestellt werden, verändern sich über die Zeit. | ① | ② | ③ | ④ | ⑤ |
| 42 | Ich verlasse mich darauf, dass ein Sachverhalt in einem pädagogischen Studientext richtig ist, wenn eine Professorin oder ein Professor diesem zustimmt. | ① | ② | ③ | ④ | ⑤ |
| 43 | Mir ist es sehr wichtig, dass sich die Inhalte von pädagogischen Studientexten für mich richtig anfühlen. | ① | ② | ③ | ④ | ⑤ |
| 44 | Was in einem pädagogischen Studientext steht, trifft in jedem Fall zu. | ① | ② | ③ | ④ | ⑤ |
| 45 | Es ist mir wichtig, dass Wissen aus pädagogischen Studientexten immer einfach anzuwenden ist. | ① | ② | ③ | ④ | ⑤ |
| 46 | Die pädagogischen Studientexte enthalten stets gesichertes Wissen. | ① | ② | ③ | ④ | ⑤ |
| 47 | Ich verlasse mich eher auf meine Intuition als auf Expertenmeinungen aus pädagogischen Studientexten. | ① | ② | ③ | ④ | ⑤ |
| 48 | Da ich mich nicht einfach auf einzelne pädagogische Studientexte verlassen kann, kommt mir bei der Konstruktion meines Wissens eine wesentliche Aufgabe zu. | ① | ② | ③ | ④ | ⑤ |
| 49 | Es ist wichtig beim Lesen eines pädagogischen Studientextes nach den spezifischen Fakten zu suchen. | ① | ② | ③ | ④ | ⑤ |
| 50 | Für die Anwendungsbereiche im Studium reicht es aus, wenn man Inhalte von pädagogischen Studientexten im Großen und Ganzen einfach mal glaubt. | ① | ② | ③ | ④ | ⑤ |
| 51 | Es ist wichtig, dass die Sachverhalte in pädagogischen Studientexten einfach dargestellt werden. | ① | ② | ③ | ④ | ⑤ |
| 52 | Es ist Aufgabe der Expertinnen und Experten, die inhaltlichen Differenzen im Wissen auszudiskutieren. Erst danach sollten sie pädagogische Studientexte verfassen. | ① | ② | ③ | ④ | ⑤ |

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

| Wie stark stimmen Sie den Aussagen zu in Bezug auf Ihren **Studienbereich Pädagogik/Schulpädagogik/Erziehungswissenschaft**? | | **trifft gar nicht zu** | **trifft eher nicht zu** | **teils teils** | **trifft eher zu** | **trifft voll zu** |
|---|---|---|---|---|---|---|
| 53 | Ich vertraue den Inhalten von pädagogischen Studientexten, auch wenn die Autorinnen und Autoren keine Begründungen und Belege anführen. | ① | ② | ③ | ④ | ⑤ |
| 54 | Die echte Welt ist viel komplexer als dies in unseren pädagogischen Studientexten oft dargestellt wird. | ① | ② | ③ | ④ | ⑤ |
| 55 | Es kommt bei einem pädagogischen Studientext darauf an, dass viele unterschiedliche Ansichten diskutiert werden. | ① | ② | ③ | ④ | ⑤ |
| 56 | Im Studium muss man sich darauf verlassen, dass das, was man in pädagogischen Studientexten liest, zutrifft. | ① | ② | ③ | ④ | ⑤ |
| 57 | Wirklich anwendbares Wissen muss in pädagogischen Studientexten eher komplex dargestellt werden. | ① | ② | ③ | ④ | ⑤ |
| 58 | Gute Theorien in pädagogischen Studientexten erkennt man daran, dass sie sich auf viele verschiedene Forschungsergebnisse stützen. | ① | ② | ③ | ④ | ⑤ |
| 59 | Trotz hohen Ansehens einer Autorin oder eines Autors, ist es wahrscheinlich, dass Falsches in den pädagogischen Studientexten steht. | ① | ② | ③ | ④ | ⑤ |
| 60 | Weil sich Wissen immer stark verändern wird, ist es völlig normal, wenn sich verschiedene pädagogische Studientexte inhaltlich widersprechen. | ① | ② | ③ | ④ | ⑤ |
| 61 | Wenn ein pädagogischer Studientext aus einem auflagenstarken Lehrbuch kommt, dann kann man sich sicher darauf verlassen, dass der Inhalt vertrauenswürdig ist. | ① | ② | ③ | ④ | ⑤ |
| 62 | In pädagogischen Studientexten werden Dinge oft unnötig verkompliziert. | ① | ② | ③ | ④ | ⑤ |
| 63 | Expertinnen und Experten schreiben nur vollständig gesicherte Erkenntnisse in pädagogische Studientexte. | ① | ② | ③ | ④ | ⑤ |
| 64 | Einen pädagogischen Studientext verstehen heißt, eher das große Ganze zu verstehen als die Details. | ① | ② | ③ | ④ | ⑤ |
| 65 | Viele Ansichten in heutigen pädagogischen Studientexten entsprechen nicht mehr dem, was früher als richtig angenommen wurde. | ① | ② | ③ | ④ | ⑤ |
| 66 | Pädagogische Studientexte, in denen die Autorin oder ein Autor nur Behauptungen aufstellt, ohne diese zu belegen oder gut zu begründen, halte ich für weniger vertrauenswürdig. | ① | ② | ③ | ④ | ⑤ |
| 67 | Es ist wichtig, dass Überschneidungen von Inhaltsbereichen in pädagogischen Studientexten vermieden werden. | ① | ② | ③ | ④ | ⑤ |
| 68 | Ich bin der Überzeugung, dass Erkenntnisse, die sich oft verändern, erst gar nicht in pädagogischen Studientexten vorkommen sollten. | ① | ② | ③ | ④ | ⑤ |
| 69 | Wenn ich die in einem pädagogischen Studientext beschriebene Position kenne, bin ich mir sicher, mich in dem Forschungsbereich gut auszukennen. | ① | ② | ③ | ④ | ⑤ |
| 70 | Wirklich relevantes Wissen ist in pädagogischen Studientexten immer widerspruchsfrei. | ① | ② | ③ | ④ | ⑤ |

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

| | Wie stark stimmen Sie den Aussagen zu in Bezug auf Ihren **Studienbereich Pädagogik/Schulpädagogik/Erziehungs-wissenschaft**? | **trifft gar nicht zu** | **trifft eher nicht zu** | **teils teils** | **trifft eher zu** | **trifft voll zu** |
|---|---|---|---|---|---|---|
| 71 | Es ist mir wichtig, dass die in pädagogischen Studientexten dargestellten Meinungen durch ausgewiesene Autoritäten legitimiert sind. | ① | ② | ③ | ④ | ⑤ |
| 72 | Beim Lesen eines pädagogischen Studientextes ist es das Wichtigste zu erkennen, welche der dargestellten Lösungen die richtige ist. | ① | ② | ③ | ④ | ⑤ |
| 73 | Heutzutage finden Forschende eher nur sehr spezielle Kleinigkeiten an neuem pädagogischen Wissen heraus, deshalb bleiben die Inhalte von Studientexten im Großen und Ganzen über die Jahrzehnte gleich. | ① | ② | ③ | ④ | ⑤ |
| 74 | Um einschätzen zu können, ob die in einem pädagogischen Studientext dargestellte Meinung verlässlich ist, ist es wichtig, dass ich versuche Verbindungen zu meinem Vorwissen herzustellen. | ① | ② | ③ | ④ | ⑤ |
| 75 | Es ist besser, sich bei pädagogischen Studientexten an die großen Autoritäten im Fach zu halten. | ① | ② | ③ | ④ | ⑤ |
| 76 | Es ist wichtig, dass Studierende über eigene Methoden der Überprüfung verfügen, um die Reichweite des Wissens in pädagogischen Studientexten einschätzen zu können. | ① | ② | ③ | ④ | ⑤ |
| 77 | Bei pädagogischen Studientexten kommen über die Jahre zwar ein paar Dinge hinzu aber dennoch bleiben die Erkenntnisse weitestgehend erhalten. | ① | ② | ③ | ④ | ⑤ |
| 78 | Nicht pädagogische Studientexte, sondern eigene Erfahrung ist der beste Weg, um etwas in diesem Fach herauszufinden. | ① | ② | ③ | ④ | ⑤ |
| 79 | Manchmal muss man die in einem pädagogischen Studientext gegebene Antwort einfach akzeptieren, auch wenn man sie nicht versteht. | ① | ② | ③ | ④ | ⑤ |
| 80 | Ich muss nachvollziehen können, wie genau eine pädagogische Erkenntnis im Studientext zustande kam, um ihr Glauben zu schenken. | ① | ② | ③ | ④ | ⑤ |
| 81 | Wenn in einem pädagogischen Studientext zwei verschiedene Meinungen zu einem Thema dargestellt werden, kann maximal nur eine davon richtig sein. | ① | ② | ③ | ④ | ⑤ |
| 82 | Pädagogische Studientexte werden von meinen Hochschul-lehrenden korrekt ausgewählt, so dass ich mich auf den Inhalt verlassen kann. | ① | ② | ③ | ④ | ⑤ |
| 83 | Autorinnen und Autoren von pädagogischen Studientexten ändern häufiger ihre Meinung darüber, was in ihrem Fach wahr ist. | ① | ② | ③ | ④ | ⑤ |
| 84 | Für die Vertrauenswürdigkeit des enthaltenen Wissens spielt es eine wichtige Rolle, wann ein pädagogischer Studientext publiziert wurde. | ① | ② | ③ | ④ | ⑤ |
| 85 | Es ist als Studierende/r wichtig, neben pädagogischen Studientexten noch weiteren Quellen zurate zu ziehen, um verschiedene Forschungsmeinungen zu konsultieren. | ① | ② | ③ | ④ | ⑤ |
| 86 | Ich erwarte von meinen pädagogischen Studientexten gesichertes Wissen. | ① | ② | ③ | ④ | ⑤ |

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

| Wie stark stimmen Sie den Aussagen zu in Bezug auf Ihren **Studienbereich Pädagogik/Schulpädagogik/Erziehungs-wissenschaft**? | | **trifft gar nicht zu** | **trifft eher nicht zu** | **teils teils** | **trifft eher zu** | **trifft voll zu** |
|---|---|---|---|---|---|---|
| 87 | Mir ist es wichtig einzuschätzen, ob das, was ich lese, in sich logisch ist, um beim Lesen eines pädagogischen Studientextes zu Erkenntnissen zu gelangen. | ① | ② | ③ | ④ | ⑤ |
| 88 | Was heute in pädagogischen Studientexten als wahr gilt, trifft in Zukunft auch noch zu. | ① | ② | ③ | ④ | ⑤ |
| 89 | Brauchbares Wissen wird in pädagogischen Studientexten immer in einfache Ursache-Wirkung-Beziehungen herunter gebrochen. | ① | ② | ③ | ④ | ⑤ |
| 90 | Es ist mir unwichtig, ob mir Inhalte von pädagogischen Studientexten sofort plausibel erscheinen. | ① | ② | ③ | ④ | ⑤ |
| 91 | Forschende entdecken sehr oft Relevantes und Neues, so dass ich mein pädagogisches Wissen in kürzerer Zeit mittels neuer Studientexte immer wieder aktualisieren sollte. | ① | ② | ③ | ④ | ⑤ |
| 92 | Meine eigenen pädagogischen Ansichten erlange ich aus sehr vielen verschiedenen Quellen. | ① | ② | ③ | ④ | ⑤ |
| 93 | Ich werde mich schon in wenigen Jahren den meisten Inhalten aus den pädagogischen Studientexten neu widmen müssen, weil viele Dinge wieder überholt sein werden. | ① | ② | ③ | ④ | ⑤ |
| 94 | Man kann in der Regel alles glauben, was man in pädagogischen Studientexten liest. | ① | ② | ③ | ④ | ⑤ |
| 95 | Ich bin der Überzeugung, dass Autorinnen und Autoren, die die Sachverhalte in pädagogischen Studientexten komplex darstellen, sich eigentlich gar nicht so gut auskennen. | ① | ② | ③ | ④ | ⑤ |
| 96 | Ich bin offen gegenüber Ansichten in pädagogischen Studientexten, die ich bisher noch nicht kannte. | ① | ② | ③ | ④ | ⑤ |
| 97 | Wenn ein pädagogischer Studientext ein bestimmtes Vorgehen empfiehlt, finde ich es wichtig, mich daran zu halten. | ① | ② | ③ | ④ | ⑤ |
| 98 | Die meisten Autorinnen und Autoren von pädagogischen Studientexten, stellen die Themen komplizierter dar, als sie eigentlich sind. | ① | ② | ③ | ④ | ⑤ |

**Haben Sie Anmerkungen?**

**Herzlichen Dank für Ihre Teilnahme!**

## StEB-Version der Studie 2 mit LIST-R

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

## Allgemeine Angaben

Bitte tragen Sie die ersten drei Buchstaben des Vornamens Ihrer Mutter, die ersten drei Buchstaben des Vornamens Ihres Vaters, Ihren Geburts**tag** sowie die ersten drei Buchstaben Ihres Geburtsortes als Code ein:

Persönlicher Code

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | |

(Vorname Mutter, z.B. KER für Kerstin) (Vorname Vater, z.B. JÖR für Jörg) (Geburts**tag**, z.B. 06 für 6. Mai 1995) (Geburtsort, z.B. MÜN für München)

Alter

Geschlecht
W ☐
M ☐

Name Ihrer Hochschule

Studiengang
Lehramt ☐
Bachelor ☐
Master ☐

Angestrebte Schulart (z.B. Grundschule)

Studienbeginn Ihres aktuellen Studiengangs
WiSe / (z.B. 2013/14)
SoSe (z.B. 2014)

Fächerkombination im Lehramt
(Bitte alle Fächer angeben.)
(z.B. Englisch, Sport)
(Bitte alle Fächer vollständig ausschreiben z.B. Sachunterricht statt SU.)

Anzahl der Fachsemester (Anzahl der Semester in Ihrem aktuellen Studienfach.)

Anzahl der Hochschulsemester (Anzahl der Semester, die Sie bereits insgesamt an einer Hochschule eingeschrieben sind.)

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

# Einstellungen zu pädagogischen Studientexten

In diesem Fragebogen geht es um Ihre Ansichten zu Studientexten im Studienbereich Pädagogik, Schulpädagogik und Erziehungswissenschaft.

**Wichtige Hinweise zum Ausfüllen des folgenden Fragebogens:**

Bitte lesen Sie sich alle Aussagen sorgfältig durch und entscheiden Sie jeweils auf einer Skala von 1 (trifft gar nicht zu) bis 5 (trifft voll zu), inwieweit Sie der Aussage zustimmen bzw. inwieweit die Aussage auf Sie zutrifft.

**Bitte beziehen Sie die Aussagen ausschließlich auf Ihre erziehungswissenschaftlichen Studienanteile (das umfasst z.B. Schulpädagogik, Allgemeine Pädagogik, Bildungswissenschaft).**

Bitte geben Sie ehrlich Ihre Meinung an. Es gibt keine falschen Antworten. Richtig ist, was Sie denken und Ihre Daten werden vollständig vertraulich behandelt.

Es ist sehr wichtig, dass Sie alle Aussagen auf dem Bogen beantworten. Bitte lassen Sie also keine Aussage weg, auch wenn Fragen vorkommen, die ähnlich klingen.

Bitte machen Sie nur ein Kreuz je Aussage.

**Wichtig!**

Unter **pädagogischen Studientexten** werden in diesem Fragebogen alle wissenschaftlichen Texte verstanden, die Sie in den erziehungswissenschaftlichen Studienbereichen, wie z.B. **Schulpädagogik, Allgemeine Pädagogik und Bildungswissenschaft**, für Ihr Studium lesen.

Das umfasst sowohl Texte, die Sie sich selbst suchen sowie solche, die Ihnen von Hochschullehrenden für Veranstaltungen (z.B. Vorlesung, Seminar, Übung) bereitgestellt werden (z.B. auf Lernplattformen wie ILIAS, Moodle oder StudIP).

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

| Wie stark stimmen Sie den Aussagen zu in Bezug auf Ihre **erziehungswissenschaftlichen Studienanteile (Schulpädagogik, Allgemeine Pädagogik, Bildungswissenschaft)**? | | **trifft gar nicht zu** | **trifft eher nicht zu** | **teils teils** | **trifft eher zu** | **trifft voll zu** |
|---|---|---|---|---|---|---|
| 1 | Ich bin offen gegenüber Ansichten in pädagogischen Studientexten, die ich bisher noch nicht kannte. | ① | ② | ③ | ④ | ⑤ |
| 2 | Da ich nicht prüfen kann, ob stimmt was ich in pädagogischen Studientexten lese, glaube ich es erstmal. | ① | ② | ③ | ④ | ⑤ |
| 3 | Es ist Aufgabe der Expertinnen und Experten, die inhaltlichen Differenzen im Wissen auszudiskutieren. Erst danach sollten sie pädagogische Studientexte verfassen. | ① | ② | ③ | ④ | ⑤ |
| 4 | Bei pädagogischen Studientexten kommen über die Jahre zwar ein paar Dinge hinzu aber dennoch bleiben die Erkenntnisse weitestgehend erhalten. | ① | ② | ③ | ④ | ⑤ |
| 5 | Es ist als Studierende/r wichtig, neben pädagogischen Studientexten noch weiteren Quellen zurate zu ziehen, um verschiedene Forschungsmeinungen zu konsultieren. | ① | ② | ③ | ④ | ⑤ |
| 6 | Es ist wichtig, dass die Sachverhalte in pädagogischen Studientexten einfach dargestellt werden. | ① | ② | ③ | ④ | ⑤ |
| 7 | Was man in einem pädagogischen Studientext liest, ist in der Regel wahr. | ① | ② | ③ | ④ | ⑤ |
| 8 | Die Vorstellungen, die in pädagogischen Studientexten dargestellt werden, verändern sich über die Zeit. | ① | ② | ③ | ④ | ⑤ |
| 9 | Ich bin der Überzeugung, dass sich wirklich nützliches Wissen in pädagogischen Studientexten in der Regel sehr einfach darstellen lässt. | ① | ② | ③ | ④ | ⑤ |
| 10 | Wenn ein pädagogischer Studientext aus einem auflagenstarken Lehrbuch kommt, dann kann man sich sicher darauf verlassen, dass der Inhalt vertrauenswürdig ist. | ① | ② | ③ | ④ | ⑤ |
| 11 | Es ist mir wichtig einen pädagogischen Studientext nicht nur zu lesen, sondern selbst zu durchdenken. | ① | ② | ③ | ④ | ⑤ |
| 12 | Ich werde mich schon in wenigen Jahren den meisten Inhalten aus den pädagogischen Studientexten neu widmen müssen, weil viele Dinge wieder überholt sein werden. | ① | ② | ③ | ④ | ⑤ |
| 13 | Pädagogische Studientexte werden von meinen Hochschullehrenden korrekt ausgewählt, so dass ich mich auf den Inhalt verlassen kann. | ① | ② | ③ | ④ | ⑤ |
| 14 | Es ist mir wichtig, dass Wissen aus pädagogischen Studientexten immer einfach anzuwenden ist. | ① | ② | ③ | ④ | ⑤ |
| 15 | Es ist wichtig, die Informationen eines pädagogischen Studientextes mit denen aus anderen pädagogischen Studientexten zu verbinden. | ① | ② | ③ | ④ | ⑤ |
| 16 | In pädagogischen Studientexten werden Dinge oft unnötig verkompliziert. | ① | ② | ③ | ④ | ⑤ |
| 17 | Es ist als Studierende/r wichtig, die für mein Studium geeigneten pädagogischen Studientextquellen auszuwählen und zu bewerten. | ① | ② | ③ | ④ | ⑤ |

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

| | Wie stark stimmen Sie den Aussagen zu in Bezug auf Ihre **erziehungswissenschaftlichen Studienanteile (Schulpädagogik, Allgemeine Pädagogik, Bildungswissenschaft)**? | trifft gar nicht zu | trifft eher nicht zu | teils teils | trifft eher zu | trifft voll zu |
|---|---|---|---|---|---|---|
| 18 | Wie etwas verstanden wird, kann sich in der pädagogischen Forschung sehr schnell ändern. Demnach ändert sich dann auch, was in Studientexten berichtet wird. | ① | ② | ③ | ④ | ⑤ |
| 19 | Ich zweifle nicht an, was in pädagogischen Studientexten steht. | ① | ② | ③ | ④ | ⑤ |
| 20 | Gute pädagogische Studientexte sind diejenigen, mit denen man die Prüfungen einfach und gut besteht. | ① | ② | ③ | ④ | ⑤ |
| 21 | Forschende entdecken sehr oft Relevantes und Neues, so dass ich mein pädagogisches Wissen in kürzerer Zeit mittels neuer Studientexte immer wieder aktualisieren sollte. | ① | ② | ③ | ④ | ⑤ |
| 22 | Es ist wichtig, dass pädagogische Studientexte neben Herleitungen und Aussagen auch Auskunft darüber geben, wie die einzelnen Aspekte genau erforscht wurden. | ① | ② | ③ | ④ | ⑤ |
| 23 | Ich stelle die Autorität der Autorin oder des Autors eines pädagogischen Studientextes nicht in Frage. | ① | ② | ③ | ④ | ⑤ |
| 24 | Viele Ansichten in heutigen pädagogischen Studientexten entsprechen nicht mehr dem, was früher als richtig angenommen wurde. | ① | ② | ③ | ④ | ⑤ |
| 25 | Für mein eigenes Wissen ist es wichtig, mir über möglichst viele verschiedene Ansichten aus verschiedenen pädagogischen Studientexten einen Überblick zu verschaffen. | ① | ② | ③ | ④ | ⑤ |
| 26 | Ich bin der Überzeugung, dass Autorinnen und Autoren, die die Sachverhalte in pädagogischen Studientexten komplex darstellen, sich eigentlich gar nicht so gut auskennen. | ① | ② | ③ | ④ | ⑤ |
| 27 | Ich verlasse mich darauf, dass in einem pädagogischen Studientext steht, was in meinem Fach genau gilt und daran halte ich mich. | ① | ② | ③ | ④ | ⑤ |
| 28 | Es ist wichtig, dass pädagogische Studientexte viele Inhaltsbereiche verbinden und diese zusammenführen. | ① | ② | ③ | ④ | ⑤ |
| 29 | Ein pädagogischer Studientext sollte die richtige Antwort, d.h. Fakten enthalten, statt viele Ansichten und Theorien zu diskutieren. | ① | ② | ③ | ④ | ⑤ |
| 30 | Weil sich Wissen immer stark verändern wird, ist es völlig normal, wenn sich verschiedene pädagogische Studientexte inhaltlich widersprechen. | ① | ② | ③ | ④ | ⑤ |
| 31 | Ich verlasse mich auf Forschungserkenntnisse in pädagogischen Studientexten, wenn zugehörige Studien genannt werden. | ① | ② | ③ | ④ | ⑤ |
| 32 | Es kommt bei einem pädagogischen Studientext darauf an, dass viele unterschiedliche Ansichten diskutiert werden. | ① | ② | ③ | ④ | ⑤ |
| 33 | Wenn die Autorin oder der Autor eines pädagogischen Studientextes eine hohe Position im Fachgebiet inne hat, dann bin ich mir sicher, dass zutrifft, was sie/er schreibt. | ① | ② | ③ | ④ | ⑤ |
| 34 | Was heute in pädagogischen Studientexten als wahr gilt, trifft in Zukunft auch noch zu. | ① | ② | ③ | ④ | ⑤ |

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

Liebe Studierende,

jeder Mensch lernt oder studiert auf eine andere Weise. Dabei führen bekanntlich viele Wege zum Ziel. Mit diesem Fragebogen können Sie – in aller Kürze – kennzeichnen, wie Sie persönlich in Ihrem Studium lernen.

Im Folgenden finden Sie hier mehrere Listen mit verschiedenen Lerntätigkeiten.
Geben Sie bitte für jede Tätigkeit an, wie häufig diese bei Ihnen vorkommt.
Sie können Ihre Antworten von 1 (sehr selten) bis 5 (sehr oft) abstufen.

Eine Anmerkung:
Bei Fragebögen dieser Art gibt es immer wieder Fragen, die recht ähnlich klingen. Bitte antworten Sie auch hier so gut Sie können. Der Sinn der „Variationen" liegt darin, möglichst viele verschiedene Aspekte eines Themenbereichs richtig zu erfassen.

**Ihre Lernzeiten:**

| | | |
|---|---|---|
| Während des Semester | | h/Woche |
| Während der Semesterferien | | h/Woche |

**Wichtig!**
**Beziehen Sie alle folgenden Aussagen ausschließlich auf Ihre erziehungswissenschaftlichen Studienanteile**
**(das umfasst z.B. Schulpädagogik, Allgemeine Pädagogik, Bildungswissenschaft).**

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

## *Aktives Lernen*

| | | sehr selten 1 | selten 2 | manch-mal 3 | oft 4 | sehr oft 5 |
|---|---|---|---|---|---|---|
| | ***Inhalte organisieren*** | | | | | |
| 1 | Ich fertige Tabellen, Diagramme oder Schaubilder an, um den Stoff besser strukturiert vorliegen zu haben. | ① | ② | ③ | ④ | ⑤ |
| 2 | Ich mache mir kurze Zusammenfassungen der wichtigsten Inhalte. | ① | ② | ③ | ④ | ⑤ |
| 3 | Ich mache mir eine Gliederung mit den wichtigsten Punkten. | ① | ② | ③ | ④ | ⑤ |
| 4 | Ich versuche, den Stoff systematisch zu ordnen. | ① | ② | ③ | ④ | ⑤ |
| 5 | Ich stelle mir aus dem Lernstoff kurze Zusammenfassungen mit den Hauptideen zusammen. | ① | ② | ③ | ④ | ⑤ |
| 6 | Für größere Stoffmengen fertige ich eine Gliederung an, die die Struktur des Stoffs am besten wiedergibt. | ① | ② | ③ | ④ | ⑤ |
| 7 | Ich stelle wichtige Fachausdrücke und Definitionen in eigenen Listen zusammen. | ① | ② | ③ | ④ | ⑤ |
| | ***Inhalte wiederholen und einprägen*** | | | | | |
| 1 | Ich präge mir den Lernstoff durch Wiederholen ein. | ① | ② | ③ | ④ | ⑤ |
| 2 | Ich lese meine Aufzeichnungen mehrmals hintereinander durch. | ① | ② | ③ | ④ | ⑤ |
| 3 | Ich lerne Schlüsselbegriffe auswendig, um mich in der Prüfung besser an wichtige Inhaltsbereiche erinnern zu können. | ① | ② | ③ | ④ | ⑤ |
| 4 | Ich lese einen Abschnitt und versuche dann, ihn auswendig wiederzugeben. | ① | ② | ③ | ④ | ⑤ |
| 5 | Ich lerne Regeln, Fachbegriffe oder Formeln auswendig. | ① | ② | ③ | ④ | ⑤ |
| 6 | Ich gehe in Gedanken noch einmal den gelernten Stoff durch, um zu sehen, ob ich alles Wesentliche behalten habe. | ① | ② | ③ | ④ | ⑤ |
| 7 | Ich lerne den Lernstoff anhand von Skripten oder anderen Aufzeichnungen möglichst auswendig. | ① | ② | ③ | ④ | ⑤ |
| | ***Inhalte verbinden und prüfen*** | | | | | |
| 1 | Zu neuen Konzepten stelle ich mir praktische Anwendungen vor. | ① | ② | ③ | ④ | ⑤ |
| 2 | Ich versuche, neue Begriffe oder Theorien auf mir bereits bekannte Begriffe und Theorien zu beziehen. | ① | ② | ③ | ④ | ⑤ |
| 3 | Ich versuche in Gedanken, das Gelernte mit dem zu verbinden, was ich schon darüber weiß. | ① | ② | ③ | ④ | ⑤ |
| 4 | Ich denke mir konkrete Beispiele zu bestimmten Lerninhalten aus. | ① | ② | ③ | ④ | ⑤ |
| 5 | Ich prüfe, ob Theorien, Interpretationen oder Schlussfolgerungen ausreichend belegt und begründet sind. | ① | ② | ③ | ④ | ⑤ |
| 6 | Ich denke über Alternativen zu den Behauptungen oder Schlussfolgerungen nach. | ① | ② | ③ | ④ | ⑤ |
| 7 | Der neue Stoff dient mir als Ausgangspunkt für die Entwicklung eigener Ideen. | ① | ② | ③ | ④ | ⑤ |

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

## *Lernen organisieren*

| | | sehr selten 1 | selten 2 | manch-mal 3 | oft 4 | sehr oft 5 |
|---|---|---|---|---|---|---|
| | ***Meine Planung*** | | | | | |
| 1 | Ich lege im vornhinein fest, wie weit ich mit der Durcharbeitung des Stoffs kommen möchte. | ① | ② | ③ | ④ | ⑤ |
| 2 | Vor dem Lernen überlege ich mir, wie ich am effektivsten vorgehen kann. | ① | ② | ③ | ④ | ⑤ |
| 3 | Ich überlege mir, in welcher Reihenfolge ich den Stoff am besten durcharbeite. | ① | ② | ③ | ④ | ⑤ |
| | ***Meine Qualitätskontrolle*** | | | | | |
| 1 | Ich stelle mir Fragen zum Stoff, um sicher zu gehen, dass ich alles verstanden habe. | ① | ② | ③ | ④ | ⑤ |
| 2 | Um Wissenslücken festzustellen, rekapituliere ich die wichtigsten Inhalte, ohne meine Unterlagen zu Hilfe zu nehmen. | ① | ② | ③ | ④ | ⑤ |
| 3 | Ich bearbeite zusätzliche Aufgaben, um festzustellen, ob ich den Stoff wirklich verstanden habe. | ① | ② | ③ | ④ | ⑤ |
| 4 | Wenn ich Schwierigkeiten beim Lernen bestimmter Inhalte habe, gehe ich meine Arbeit noch einmal langsam durch. | ① | ② | ③ | ④ | ⑤ |
| 5 | Wenn ich während des Lesens nicht alles verstehe, versuche ich die Lücken festzuhalten und den Text daraufhin noch einmal durchzugehen. | ① | ② | ③ | ④ | ⑤ |
| 6 | Wenn mir eine bestimmte Textstelle unklar erscheint, versuche ich durch genaues Lesen die Unklarheit zu beseitigen. | ① | ② | ③ | ④ | ⑤ |
| 7 | Wenn ich einen Sachverhalt beim ersten Mal nicht verstehe, gehe ich ihn noch einmal Schritt für Schritt durch. | ① | ② | ③ | ④ | ⑤ |

**Wichtig!**
**Beziehen Sie auch die folgenden Aussagen ausschließlich auf Ihre erziehungswissenschaftlichen Studienanteile**
**(das umfasst z.B. Schulpädagogik, Allgemeine Pädagogik, Bildungswissenschaft).**

Lehrstuhl für Erziehungswissenschaft
mit Schwerpunkt Empirische Lehr-Lernforschung

UNIVERSITÄT PASSAU

## *Ressourcen nutzen*

| | | sehr selten 1 | selten 2 | manch-mal 3 | oft 4 | sehr oft 5 |
|---|---|---|---|---|---|---|
| | ***Lernzeit*** | | | | | |
| 1 | Beim Lernen halte ich mich an einen bestimmten Zeitplan. | ① | ② | ③ | ④ | ⑤ |
| 2 | Ich lege bestimmte Zeiten fest, zu denen ich dann lerne. | ① | ② | ③ | ④ | ⑤ |
| 3 | Ich lege die Stunden, die ich täglich mit Lernen verbringe, durch einen Zeitplan fest. | ① | ② | ③ | ④ | ⑤ |
| 4 | Ich lege vor jeder Lernphase eine bestimmte Zeitdauer fest. | ① | ② | ③ | ④ | ⑤ |
| | ***Lernumgebung*** | | | | | |
| 1 | Ich lerne an einem Platz, an dem ich mich wohlfühle. | ① | ② | ③ | ④ | ⑤ |
| 2 | Ich lerne an einem Platz, an dem ich konzentriert arbeiten kann. | ① | ② | ③ | ④ | ⑤ |
| 3 | Ich gestalte meine Umgebung so, dass ich möglichst nicht abgelenkt werde. | ① | ② | ③ | ④ | ⑤ |
| 4 | Wenn ich lerne, sorge ich dafür, dass ich in Ruhe arbeiten kann. | ① | ② | ③ | ④ | ⑤ |
| 5 | Mein Arbeitsplatz ist so gestaltet, dass ich alles schnell finden kann. | ① | ② | ③ | ④ | ⑤ |
| 6 | Ich habe an meinem Arbeitsplatz die wichtigsten Unterlagen griffbereit. | ① | ② | ③ | ④ | ⑤ |
| | ***Literatur nutzen*** | | | | | |
| 1 | Wenn mir ein Inhalt noch nicht ganz klar ist, suche ich nach ergänzender Literatur. | ① | ② | ③ | ④ | ⑤ |
| 2 | Wenn ich einen Begriff nicht verstehe, schlage ich in der Literatur nach. | ① | ② | ③ | ④ | ⑤ |
| 3 | Fehlende Informationen suche ich mir aus verschiedenen Quellen zusammen (z.B. Bücher, Fachzeitschriften). | ① | ② | ③ | ④ | ⑤ |
| 4 | Bei Bedarf ziehe ich zusätzliche Literatur heran. | ① | ② | ③ | ④ | ⑤ |
| | ***Online-Quellen nutzen*** | | | | | |
| 1 | Wenn ich zusätzliche Literatur suche, benutze ich entsprechende Datenbanken im Netz. | ① | ② | ③ | ④ | ⑤ |
| 2 | Ich nutze für mein Studium Online Lexika (z.B. Wikipedia). | ① | ② | ③ | ④ | ⑤ |
| 3 | Ich benutze bei Bedarf Suchmaschinen (Google, Yahoo, etc.). | ① | ② | ③ | ④ | ⑤ |
| 4 | Ich nutze Infos aus speziellen Internet-Foren. | ① | ② | ③ | ④ | ⑤ |
| | ***Gemeinsam lernen*** | | | | | |
| 1 | Ich lerne gemeinsam mit Studienkolleg(inn)en. | ① | ② | ③ | ④ | ⑤ |
| 2 | Ich nehme mir gerne Zeit, um mit Studienkolleg(inn)en zu diskutieren. | ① | ② | ③ | ④ | ⑤ |
| 3 | Wenn ich Verständnisprobleme habe, nehme ich die Hilfe von Studienkolleg(inn)en in Anspruch. | ① | ② | ③ | ④ | ⑤ |
| 4 | Wenn mir etwas unklar ist, wende ich mich an meine Studienkolleg(inn)en. | ① | ② | ③ | ④ | ⑤ |

**Vielen Dank für Ihre Teilnahme!**

# B Itemwortlaut der StEB-Versionen

## Itempool mit 126 Items

TABELLE B.1: Itemwortlaut, -polung und -ursprung für den Itempool – Teil 1

| Item-code | Itemwortlaut | Pol-ung |
|---|---|---|
| sw_1 | Die Dinge sind einfacher als sie in den meisten Studientexten dargestellt werden. (S/16) | 1 |
| sw_2 | Wenn es nicht gelingt, etwas in Studientexten einfach dazustellen, dann ist es wahrscheinlich falsch. | 1 |
| sw_3 | Wenn Autorinnen und Autoren etwas komplex darstellen, dann versuchen sie damit oft zu verbergen, dass sie sich eigentlich gar nicht so gut auskennen. | 1 |
| sw_4 | Wenn etwas komplex ist, dann führt auch kein Weg daran vorbei, diesen Sachverhalt auch in seiner ganzen Komplexität im Studientext darzustellen. | 0 |
| sw_5 | Studientexte dürfen auf keinen Fall widersprüchlich sein. | 1 |
| sw_6 | Bei gehäuften Widersprüchen in Studientexten ist eigentlich immer etwas faul. | 1 |
| sw_7 | Bei guten Studientexten wird streng nach einzelnen Inhalten getrennt. | 1 |
| sw_8 | Bei guten Studientexten wird vermieden, dass sich die verschiedenen Bereiche zu sehr überlappen. | 1 |
| sw_9 | Gute Studientexte verbinden viele Inhaltsbereiche und führen diese zusammen. | 0 |
| sw_10 | Gute Studientexte gehen auf die verschiedenartigsten Verbindungen zu inhaltlichen Nachbarfeldern ein. | 0 |
| sw_11 | Nützliches Wissen lässt sich in Studientexten immer auch sehr einfach darstellen. | 1 |
| sw_12 | Sobald Wissen aus Studientexten nützlich wird, ist es eigentlich auch immer recht komplex. | 0 |

S: Item basiert auf Item von Schommer (1998, 561 f.) / Itemnummer

TABELLE B.2: Itemwortlaut, -polung und -ursprung für den Itempool – Teil 2

| Item-code | Itemwortlaut | Pol-ung |
|---|---|---|
| sw_13 | Nützliches Studientextwissen ist wie ein Rezept organisiert: Es werden in der Reihenfolge der Handlungen alle wesentlichen Schritte genannt. | 1 |
| sw_14 | Weil die echte Welt so vielfältig ist, ist auch das Wissen darüber komplex und sogar oft widersprüchlich. | 0 |
| sw_15 | In Studientexten werden Dinge oft unnötig verkompliziert. | 1 |
| sw_16 | Die echte Welt ist viel komplexer als das in unseren Studientexten oft dargestellt wird. | 0 |
| sw_17 | Wissen aus Studientexten muss immer einfach anzuwenden sein. | 1 |
| sw_18 | Wirklich wirksames Wissen muss eher komplex dargestellt werden. | 0 |
| sw_19 | Echtes Wissen ist auch in Studientexten immer widerspruchsfrei. | 1 |
| sw_20 | Brauchbares Wissen wird in Studientexten immer in einfache Ursache-Wirkung-Beziehungen herunter gebrochen. | 1 |
| sw_21 | Wenn man sich die Fakten aus den Studientexten aneignet, weiß man, was es zu wissen gibt. | 1 |
| sw_22 | Ich versuche die Informationen eines Studientextes mit denen aus anderen Studientexten zu verbinden. (S/14) | 0 |
| sw_23 | Beim Lesen eines Studientextes suche ich nach den spezifischen Fakten. (S/38) | 1 |
| sw_24 | Einen Studientext verstehen heißt, eher das große Ganze zu verstehen, als die Details. (S/18) | 0 |
| sw_25 | Ein Studientext ist gut, wenn er die wichtigsten Fakten zum Thema enthält. | 0 |
| sw_26 | Ein Studientext sollte alle Fakten enthalten, die zu dem Themenbereich wichtig sind. | 1 |

S: Item basiert auf Item von Schommer (1998, 561 f.) / Itemnummer

TABELLE B.3: Itemwortlaut, -polung und -ursprung für den Itempool – Teil 3

| Item-code | Itemwortlaut | Pol-ung |
|---|---|---|
| qw_1 | Was in einem Studientext steht, kann man auch genau so einsetzen. | 1 |
| qw_2 | Expertinnen und Experten schreiben nur vollständig gesicherte Erkenntnisse in Studientexte. | 1 |
| qw_3 | Bevor etwas in einem Studientext landet, wird sichergestellt, dass das Wissen auch wahr ist. | 1 |
| qw_4 | Expertinnen und Experten unterliegen strengen Regeln, die für uns sicherstellen, dass das Wissen in Studientexten weitgehend wahr ist. | 1 |
| qw_5 | Auch die besten Expertinnen und Experten irren sich häufiger und müssen ihre Ansichten später korrigieren. | 0 |
| qw_6 | Es ist fast egal, wie angesehen eine Autorin oder ein Autor ist. Es kann trotzdem Falsches in den Studientexten stehen. | 0 |
| qw_7 | Lesende von Studientexten müssen über eigene Methoden der Überprüfung verfügen, um die Reichweite des Wissens darin einschätzen zu können. | 0 |
| qw_8 | Für die Anwendungsbereiche im Studium reicht es aus, wenn man Inhalte von Studientexten im Großen und Ganzen einfach mal glaubt. | 1 |
| qw_9 | Angesehene Studientextautorinnen und -autoren gewinnen ihr Ansehen dadurch, dass sie viel öfter richtig liegen als andere. | 1 |
| qw_10 | Es ist besser, sich bei Studientexten an die großen Autoritäten im Fach zu halten. | 1 |
| qw_11 | Für mein eigenes Wissen ist es wichtig, mir über möglichst vielen verschiedenen Ansichten einen Überblick zu verschaffen. | 0 |
| qw_12 | Meine eigenen Ansichten erlange ich aus sehr vielen verschiedenen Quellen. | 0 |
| qw_13 | Ich bin darauf angewiesen, dass mir jemand sagt, was genau gilt, damit ich mich dann später daran halten kann. | 1 |
| qw_14 | Wenn es in einem Bereich keine anerkannten Expertinnen und Experten gibt, dann muss ich mich eher auf meine Intuition und Erfahrung verlassen. | 1 |
| qw_15 | Meine eigene Erfahrung vermag mehr Wahrheit zu erkennen als die meisten theoretischen Texte. | 1 |
| qw_16 | Für die Vertrauenswürdigkeit des enthaltenen Wissens spielt es eine wichtige Rolle, wann ein Studientext publiziert wurde. | 0 |
| qw_17 | Gute Studientexte sind diejenigen, mit denen man die Prüfungen einfach und gut besteht. Sie enthalten das richtige Wissen. | 1 |
| qw_18 | Wenn ich in einer Prüfung eine sehr gute Note bekomme, dann habe ich auch genau die wichtigen Sachen aus den Studientexten gelernt. | 1 |
| qw_19 | Wenn eine Professorin oder ein Professor einer Studientextmeinung zustimmt, dann ist sie wahrscheinlich richtig. | 1 |
| qw_20 | Was in Studientexten steht, muss man glauben. (C/2, U/2) | 1 |
| qw_21 | Was man in einem Studientext liest, ist sicher wahr. (H/1, C/4, U/4) | 1 |
| qw_22 | Nur die Autorinnen und Autoren der Studientexte wissen genau, was in ihrem Fach zutrifft. (C/5) | 1 |

C: Item basiert auf Item von Conley et al. (2004, S. 202 f.) / Itemnummer
H: Item basiert auf Item von Hofer (2000, 390 f.) / Itemnummer
U: Item basiert auf Item von Urhahne & Hopf (2004, S. 84 f.) / Itemnummer

TABELLE B.4: Itemwortlaut, -polung und -ursprung für den Itempool – Teil 4

| Item-code | Itemwortlaut | Pol-ung |
|---|---|---|
| qw_23 | Im Studium muss man sich darauf verlassen, dass das, was man in Studientexten liest, zutrifft. | 1 |
| qw_24 | Neues Wissen erlangt man, indem man es sich u.a. unter Zuhilfenahme von Studientexten konstruiert. | 0 |
| qw_25 | Studierende sollten nicht anzweifeln, was in Studientexten steht. | 1 |
| qw_26 | Studierende sollten die Autorität der Autorin oder des Autors eines Studientextes nicht in Frage stellen. (SBD/26) | 1 |
| qw_27 | Wenn ein Studientext ein bestimmtes Vorgehen empfiehlt, halte ich mich daran. (SBD/25) | 1 |
| qw_28 | Man kann fast alles glauben, was man in Studientexten liest. (S/6) | 1 |
| qw_29 | Wenn mir beim Lesen eines Studientextes auffällt, dass dieser meiner Erfahrung nicht entspricht, dann muss ich mich wohl geirrt haben. (H/2) | 1 |
| qw_30 | Ich bin mir sehr sicher etwas zu wissen, wenn ich die in einem Studientext beschriebene Position kenne. (H/3) | 1 |
| qw_31 | Manchmal muss man die in einem Studientext gegebene Antwort einfach akzeptieren, auch wenn man sie nicht versteht. (H/4) | 1 |
| rw_1 | In Studientexten dargestelltes Wissen muss mir in erster Linie plausibel sein. | 1 |
| rw_2 | Mir ist es sehr wichtig, dass sich die Inhalte von Studientexten für mich richtig anfühlen. | 1 |
| rw_3 | Es ist weniger wichtig ob mir etwas in Studientexten sofort plausibel erscheint. | 0 |
| rw_4 | Es ist mir wichtig, dass die Studientextmeinung durch ausgewiesene Autoritäten legitimiert ist. | 1 |
| rw_5 | Ich muss nachvollziehen können, wie genau eine Erkenntnis im Studientext zustande kam, um ihr Glauben zu schenken. | 0 |
| rw_6 | Mit den richtigen Methoden kann auch jemand, der keine große Autorität ist, zu bahnbrechenden Erkenntnissen gelangen. | 0 |
| rw_7 | Für mich ist es schon wichtig, wer etwas geschrieben hat, wenn ich dem Inhalt Glauben schenken soll. | 1 |
| rw_8 | Studientexte werden von meinen Hochschullehrenden korrekt ausgewählt, so dass ich mich auf sie verlassen kann. | 1 |
| rw_9 | Es ist Aufgabe meiner Hochschullehrenden, die richtigen Studientexte für mich auszuwählen. | 1 |
| rw_10 | Ich suche oft nach weiteren Quellen, weil mir die von meinen Hochschullehrenden ausgewählten Studientexte nicht ganz ausreichen. | 0 |
| rw_11 | Ich ziehe neben Studientexten noch weiteren Quellen zurate, um verschiedene Forschungsmeinungen zu konsultieren. | 0 |
| rw_12 | Es ist als Studierende/r eher meine Aufgabe, die für mein Studium geeigneten Studientextquellen auszuwählen und zu bewerten. | 0 |
| rw_13 | Mir selbst kommt bei der Gestaltung meines Wissens eine wesentliche Aufgabe zu, weil ich mich nicht einfach auf einzelne Studientexte verlassen kann. | 0 |

H: Item basiert auf Item von Hofer (2000, 390 f.) / Itemnummer
SDB: Item basiert auf Item von Schraw et al. (2002, S. 275) / Itemnummer
S: Item basiert auf Item von Schommer (1998, 561 f.) / Itemnummer

TABELLE B.5: Itemwortlaut, -polung und -ursprung für den Itempool – Teil 5

| Item-code | Itemwortlaut | Pol-ung |
|---|---|---|
| rw_14 | Die Auswahl und Gestaltung meines zu erwerbenden Wissens ist Aufgabe und Pflicht der Hochschullehrenden. | 0 |
| rw_15 | Für meinen Wissenserwerb ist es notwendig, dass meine Hochschullehrenden durch Qualifikation ausgewiesene Expertinnen und Experten auf ihrem Fachgebiet sind. | 1 |
| rw_16 | Hausarbeiten meiner Mitstudierenden sind wegen der geringeren Qualifikation in der Regel schlechter als die von Professorinnen und Professoren verfassten Lehrwerke. | 1 |
| rw_17 | Ein Studientext mit einem guten Inhalt wird üblicherweise auch von einem guten Verlag publiziert. | 1 |
| rw_18 | Ein Studientext aus einem auflagenstarken Lehrbuch spricht für einen vertrauenswürdigen Inhalt. | 1 |
| rw_19 | Wichtige Inhalte von Studientexten sind in der Regel auch spontan plausibel. | 1 |
| rw_20 | Mit genügend Zeit und Ressourcen könnte ich die meisten Forschungen und Studien aus meinen Studientexten auch selbst durchführen. | 0 |
| rw_21 | Gute Ideen können von jedem kommen, nicht nur von Autorinnen und Autoren der Studientexte. (C/20, U/22) | 0 |
| rw_22 | Gute Theorien in Studientexten erkennt man daran, dass sie sich auf viele verschiedene Forschungsergebnisse stützen. (C/22, U/24) | 0 |
| rw_23 | Um einschätzen zu können, ob die in einem Studientext dargestellte Meinung verlässlich ist, versuche ich Verbindungen zu meinem Vorwissen herzustellen. | 0 |
| rw_24 | Wenn ich einen Studientext lese, versuche ich, mir ein eigenes Erklärungsmodell zu bauen. | 0 |
| rw_25 | Um beim Lesen eines Studientextes zu echten Erkenntnissen zu gelangen, versuche ich einzuschätzen, ob das was ich lese, in sich logisch ist. | 0 |
| rw_26 | Den in einem Studientext aufgestellten Behauptungen, kann man erst trauen, wenn man sich mit mehreren Wissensquellen zu dem Thema vertraut gemacht hat. | 0 |
| rw_27 | Ich verstehe einen Studientext besser, wenn ich ihn nicht nur lese, sondern selbst durchdenke. | 0 |
| rw_28 | Ich bilde mir meine Meinung zu Themen nicht auf Basis eines einzelnen Studientextes. | 0 |
| rw_29 | Wenn die Autorin oder der Autor eines Studientextes eine hohe Position im Fachgebiet inne hat, dann wird schon zutreffen, was sie/er schreibt. | 1 |
| rw_30 | Ich vertraue den Inhalten von Studientexten, auch wenn die Autorinnen und Autoren nicht Begründungen und Belege anführen. | 1 |
| rw_31 | Wenn eine Autorin oder ein Autor nur Behauptungen aufstellt, ohne diese zu belegen oder gut zu begründen, dann halte ich den Studientext für weniger vertrauenswürdig. | 0 |

C: Item basiert auf Item von Conley et al. (2004, S. 202 f.) / Itemnummer
U: Item basiert auf Item von Urhahne & Hopf (2004, S. 84 f.) / Itemnummer

TABELLE B.6: Itemwortlaut, -polung und -ursprung für den Itempool – Teil 6

| Item-code | Itemwortlaut | Pol-ung |
|---|---|---|
| rw_32 | Nicht Studientexte, sondern eigene Erfahrung ist der beste Weg, um etwas in diesem Fach herauszufinden. (H/13) | 1 |
| rw_33 | Da ich über keine Möglichkeit verfüge zu prüfen, ob stimmt was ich in den Studientexten lese, kann ich es genauso gut einfach glauben. | 1 |
| rw_34 | Ich akzeptiere eher die Ansichten von Praktikerinnen und Praktikern des Fachs als die Ansichten von den Forschenden , die die Studientexte verfassen. (H/14) | 1 |
| gw_1 | Die Studientexte enthalten stets gesichertes Wissen. | 1 |
| gw_2 | Forschende sollten sich erstmal einig über ihre Aussagen werden, bevor sie diese in Studientexten anführen. | 1 |
| gw_3 | Es ist Aufgabe der Expertinnen und Experten, die sachlichen Differenzen im Wissen auszudiskutieren. Erst danach sollten sie Studientexte verfassen. | 1 |
| gw_4 | Sind Forschungserkenntnisse erst einmal durch Studien bewiesen, kann man sich auch darauf verlassen. | 1 |
| gw_5 | In meinem Studium bin ich darauf angewiesen, dass mir gesicherte Erkenntnisse in Studientexten vorgestellt werden. | 1 |
| gw_6 | Ich kann nicht alles überprüfen, deswegen muss ich eben glauben, was in den Studientexten steht. | 1 |
| gw_7 | Ich erwarte von meinen Studientexten gesichertes Wissen. | 1 |
| gw_8 | Studientexte sollten immer auch ungesichertes Wissen enthalten. | 0 |
| gw_9 | Studientexte sollten neben Herleitungen und Aussagen auch Auskunft darüber geben, wie die einzelnen Aspekte genau erforscht wurden. | 0 |
| gw_10 | Mir ist nicht so wichtig, wie man genau zu den Inhalten in den Studientexten kam. Wichtig ist eher, was dabei heraus kommt und ob ich das anwenden kann. | 1 |
| gw_11 | Beim Lesen eines Studientextes ist es das Wichtigste zu erkennen, welche der dargestellten Lösungen die richtige ist. (C/7, U/7) | 1 |
| gw_12 | Wenn in einem Studientext zwei verschiedene Meinungen zu einem Thema dargestellt werden, kann maximal nur eine davon richtig sein. (SDB/18) | 1 |
| gw_13 | Ein Studientext sollte die richtige Antwort enthalten, statt viele Ansichten zu diskutieren. | 1 |
| gw_14 | Die meisten Autorinnen und Autoren stellen die Themen komplizierter dar, als sie eigentlich sind. (SDB/17) | 1 |
| gw_15 | Was in einem Studientext steht, trifft in jedem Fall zu. | 1 |
| gw_16 | Es kommt bei einem Studientext darauf an, dass viele unterschiedliche Ansichten diskutiert werden. | 0 |
| gw_17 | Wenn in einem Studientext eine Ansicht vertreten wird, die ich bisher noch nicht kannte, dann bin ich dafür offen. | 1 |
| gw_18 | Wenn in einem Studientext mehrere unterschiedliche Ansichten beschrieben werden, damm kann je nach Kontext variieren, welche Sichtweise gerade die bessere ist. | 0 |

C: Item basiert auf Item von Conley et al. (2004, S. 202 f.) / Itemnummer
H: Item basiert auf Item von Hofer (2000, 390 f.) / Itemnummer
SDB: Item basiert auf Item von Schraw et al. (2002, S. 275) / Itemnummer
U: Item basiert auf Item von Urhahne & Hopf (2004, S. 84 f.) / Itemnummer

TABELLE B.7: Itemwortlaut, -polung und -ursprung für den Itempool – Teil 7

| Item-code | Itemwortlaut | Pol-ung |
|---|---|---|
| ew_1 | Die wichtigsten Anwendungsaspekte des Wissens ändern sich eigentlich nicht sonderlich über die Zeit. | 1 |
| ew_2 | Es ist völlig normal, dass Expertinnen und Experten sich über viele Bereiche uneins sind: Wissen wird sich immer stark verändern. | 0 |
| ew_3 | Wie etwas verstanden wird, kann sich in der Forschung sehr schnell ändern. | 0 |
| ew_4 | Es ist wahrscheinlich, dass der Großteil der Erkenntnisse in den Studientexten in zehn Jahren ganz anders aufgefasst wird. | 0 |
| ew_5 | Bei Studientexten kommen über die Jahre zwar ein paar Dinge hinzu aber dennoch bleiben die Erkenntnisse weitestgehend erhalten. | 1 |
| ew_6 | Es reicht aus, dass ich mich jetzt mit den Studientexten auseinandersetze. Abgesehen von Kleinigkeiten, werde ich das meiste daraus auch noch in zehn Jahren einsetzen können. | 1 |
| ew_7 | In unseren Fachbereichen reicht es aus, die Inhalte einmal gut verstanden zu haben. Es ist nicht notwendig, sich alle paar Jahre nochmal komplett neu darin zu orientieren. | 1 |
| ew_8 | Ich werde mich schon in wenigen Jahren den meisten Inhalten aus den Studientexten neu widmen müssen, weil viele Dinge wieder überholt sein werden. | 0 |
| ew_9 | Heutzutage finden Forschende eher nur sehr spezielle Kleinigkeiten an neuem Wissen heraus. Im Großen und Ganzen bleiben die Inhalte aber über die Jahrzehnte in etwa gleich. | 1 |
| ew_10 | Forschende entdecken sehr oft Relevantes und Neues, so dass ich mein Wissen in kürzerer Zeit immer wieder aktualisieren muss. | 0 |
| ew_11 | Wenn sich Erkenntnisse in einem Inhaltsbereich oft verändern, sollten sie nicht in Studientexten vorkommen. | 1 |
| ew_12 | Einige Ansichten in Studientexten entsprechen nicht mehr dem, was früher als richtig angenommen wurde. (C/10, U/12) | 0 |
| ew_13 | Die Vorstellungen, die in Studientexten dargestellt werden, verändern sich über die Zeit. (C/11, U/13) | 0 |
| ew_14 | Es gibt manche Fragen, die auch in Studientexten nicht beantwortet werden. (C/12, U/14) | 0 |
| ew_15 | Manchmal ändern Autorinnen und Autoren ihre Meinung darüber, was in ihrem Fach wahr ist. (C/15, U/17) | 0 |
| ew_16 | Was heute in Studientexten als wahr gilt, trifft in Zukunft auch noch zu. (SDB/23, H/12) | 0 |
| ew_17 | Durch neue Entdeckungen kann sich verändern, was die Autorinnen und Autoren von Studientexten für wahr halten. (U/16) | 0 |

C: Item basiert auf Item von Conley et al. (2004, S. 202 f.) / Itemnummer
H: Item basiert auf Item von Hofer (2000, 390 f.) / Itemnummer
SDB: Item basiert auf Item von Schraw et al. (2002, S. 275) / Itemnummer
U: Item basiert auf Item von Urhahne & Hopf (2004, S. 84 f.) / Itemnummer

## 98-Item StEB-Version

TABELLE B.8: Itemwortlaut, -polung und -ursprung der 98-Item StEB-Version – Teil 1

| Item-code | Itemwortlaut | Pol-ung |
|---|---|---|
| ew_2 | Weil sich Wissen immer stark verändern wird, ist es völlig normal, wenn sich verschiedene pädagogische Studientexte inhaltlich widersprechen. | 0 |
| ew_3 | Wie etwas verstanden wird, kann sich in der pädagogischen Forschung sehr schnell ändern. Demnach ändert sich dann auch, was in Studientexten berichtet wird. | 0 |
| ew_4 | Es ist wahrscheinlich, dass der Großteil der Erkenntnisse in den pädagogischen Studientexten in zehn Jahren ganz anders aufgefasst wird. | 0 |
| ew_5 | Bei pädagogischen Studientexten kommen über die Jahre zwar ein paar Dinge hinzu aber dennoch bleiben die Erkenntnisse weitestgehend erhalten. | 1 |
| ew_8 | Ich werde mich schon in wenigen Jahren den meisten Inhalten aus den pädagogischen Studientexten neu widmen müssen, weil viele Dinge wieder überholt sein werden. | 0 |
| ew_9 | Heutzutage finden Forschende eher nur sehr spezielle Kleinigkeiten an neuem pädagogischen Wissen heraus, deshalb bleiben die Inhalte von Studientexten im Großen und Ganzen über die Jahrzehnte gleich. | 1 |
| ew_10 | Forschende entdecken sehr oft Relevantes und Neues, so dass ich mein pädagogisches Wissen in kürzerer Zeit mittels neuer Studientexte immer wieder aktualisieren sollte. | 0 |
| ew_11 | Ich bin der Überzeugung, dass Erkenntnisse, die sich oft verändern, erst gar nicht in pädagogischen Studientexten vorkommen sollten. | 1 |
| ew_12 | Viele Ansichten in heutigen pädagogischen Studientexten entsprechen nicht mehr dem, was früher als richtig angenommen wurde. (C/10, U/12) | 0 |
| ew_13 | Die Vorstellungen, die in pädagogischen Studientexten dargestellt werden, verändern sich über die Zeit. (C/11, U/13) | 0 |
| ew_15 | Autorinnen und Autoren von pädagogischen Studientexten ändern häufiger ihre Meinung darüber, was in ihrem Fach wahr ist. (C/15, U/17) | 0 |
| ew_16 | Was heute in pädagogischen Studientexten als wahr gilt, trifft in Zukunft auch noch zu. (SBD/23, H/12) | 1 |

C: Item basiert auf Item von Conley et al. (2004, S. 202 f.) / Itemnummer
H: Item basiert auf Item von Hofer (2000, 390 f.) / Itemnummer
SDB: Item basiert auf Item von Schraw et al. (2002, S. 275) / Itemnummer
U: Item basiert auf Item von Urhahne & Hopf (2004, S. 84 f.) / Itemnummer
S: Item basiert auf Item von Schommer (1998, 561 f.) / Itemnummer

TABELLE B.9: Itemwortlaut, -polung und -ursprung der 98-Item StEB-Version – Teil 2

| Item-code | Itemwortlaut | Pol-ung |
|---|---|---|
| gw_1 | Die pädagogischen Studientexte enthalten stets gesichertes Wissen. | 1 |
| gw_2 | Forschende sollten sich erstmal einig über ihre Aussagen werden, bevor sie diese in pädagfogischen Studientexten anführen. | 1 |
| gw_3 | Es ist Aufgabe der Expertinnen und Experten, die inhaltlichen Differenzen im Wissen auszudiskutieren. Erst danach sollten sie pädagogische Studientexte verfassen. | 1 |
| gw_4 | Ich verlasse mich auf Forschungserkenntnisse in pädagogischen Studientexten, wenn zugehörige Studien genannt werden. | 1 |
| gw_7 | Ich erwarte von meinen pädagogischen Studientexten gesichertes Wissen. | 1 |
| gw_9 | Es ist wichtig, dass pädagogische Studientexte neben Herleitungen und Aussagen auch Auskunft darüber geben, wie die einzelnen Aspekte genau erforscht wurden. | 0 |
| gw_10 | Es ist nicht so wichtig, wie man genau zu den Inhalten in den pädagogischen Studientexten kam. Wichtig ist eher, was dabei heraus kommt und ob ich das anwenden kann. | 1 |
| gw_11 | Beim Lesen eines pädagogischen Studientextes ist es das Wichtigste zu erkennen, welche der dargestellten Lösungen die richtige ist. (C/7, U/7) | 1 |
| gw_12 | Wenn in einem pädagogischen Studientext zwei verschiedene Meinungen zu einem Thema dargestellt werden, kann maximal nur eine davon richtig sein. (SBD/18) | 1 |
| gw_13 | Ein pädagogischer Studientext sollte die richtige Antwort, d.h. Fakten enthalten, statt viele Ansichten und Theorien zu diskutieren. | 1 |
| gw_15 | Was in einem pädagogischen Studientext steht, trifft in jedem Fall zu. | 1 |
| gw_16 | Es kommt bei einem pädagogischen Studientext darauf an, dass viele unterschiedliche Ansichten diskutiert werden. | 0 |
| gw_17 | Ich bin offen gegenüber Ansichten in pädagogischen Studientexten, die ich bisher noch nicht kannte. | 0 |
| gw_18 | Wenn in einem pädagogischen Studientext mehrere unterschiedliche Ansichten beschrieben werden, dann kann je nach Kontext variieren, welche Sichtweise gerade die bessere ist. | 0 |
| gw_19* | Expertinnen und Experten schreiben nur vollständig gesicherte Erkenntnisse in pädagogische Studientexte. | 1 |
| gw_20* | Bevor etwas in einem pädagogischen Studientext landet, wird sichergestellt, dass das Wissen auch wahr ist. | 1 |

* Veränderter Itemcode aufgrund von veränderter Zuordnung des Items zu anderer Subskala: gw_19 vor Expertenbefragung ursprünglich qw_2; gw_20 vor Expertenbefragung ursprünglich qw_3.

C: Item basiert auf Item von Conley et al. (2004, S. 202 f.) / Itemnummer

SDB: Item basiert auf Item von Schraw et al. (2002, S. 275) / Itemnummer

U: Item basiert auf Item von Urhahne & Hopf (2004, S. 84 f.) / Itemnummer

TABELLE B.10: Itemwortlaut, -polung und -ursprung der 98-Item StEB-Version – Teil 3

| Item-code | Itemwortlaut | Pol-ung |
|---|---|---|
| qw_5 | Auch die besten pädagogischen Studientextautorinnen und -autoren irren sich häufiger. | 0 |
| qw_6 | Trotz hohen Ansehens einer Autorin oder eines Autors, ist es wahrscheinlich, dass Falsches in den pädagogischen Studientexten steht. | 0 |
| qw_8 | Für die Anwendungsbereiche im Studium reicht es aus, wenn man Inhalte von pädagogischen Studientexten im Großen und Ganzen einfach mal glaubt. | 1 |
| qw_10 | Es ist besser, sich bei pädagogischen Studientexten an die großen Autoritäten im Fach zu halten. | 1 |
| qw_11 | Für mein eigenes Wissen ist es wichtig, mir über möglichst viele verschiedene Ansichten aus verschiedenen pädagogischen Studientexten einen Überblick zu verschaffen. | 0 |
| qw_12 | Meine eigenen pädagogischen Ansichten erlange ich aus sehr vielen verschiedenen Quellen. | 0 |
| qw_13 | Ich verlasse mich darauf, dass in einem pädagogischen Studientext steht, was in meinem Fach genau gilt und daran halte ich mich. | 1 |
| qw_14 | Ich verlasse mich eher auf meine Intuition als auf Expertenmeinungen aus pädagogischen Studientexten. | 1 |
| qw_15 | Ich verlasse mich eher auf meine eigenen Erfahrungen als auf Expertenmeinungen aus pädagogischen Studientexten. | 1 |
| qw_19 | Ich verlasse mich darauf, dass ein Sachverhalt in einem pädagogischen Studientext richtig ist, wenn eine Professorin oder ein Professor diesem zustimmt. | 1 |
| qw_20 | Ich glaube generell schon, was in pädagogischen Studientexten steht. (C/2, U/2) | 1 |
| qw_21 | Was man in einem pädagogischen Studientext liest, ist in der Regel wahr. (H/1, C/4, U/4) | 1 |
| qw_22 | Nur die Autorinnen und Autoren der pädagogischen Studientexte wissen genau, was in ihrem Fach zutrifft. (C/5) | 1 |
| qw_23 | Im Studium muss man sich darauf verlassen, dass das, was man in pädagogischen Studientexten liest, zutrifft. | 1 |
| qw_24 | Neues Wissen erlangt man in der Pädagogik, indem man es sich selbst u.a. unter Zuhilfenahme von Studientexten konstruiert. | 0 |
| qw_25 | Ich zweifle nicht an, was in pädagogischen Studientexten steht. | 1 |
| qw_26 | Ich stelle die Autorität der Autorin oder des Autors eines pädagogischen Studientextes nicht in Frage. (SBD/26) | 1 |
| qw_27 | Wenn ein pädagogischer Studientext ein bestimmtes Vorgehen empfiehlt, finde ich es wichtig, mich daran zu halten. (SBD/25) | 1 |
| qw_28 | Man kann in der Regel alles glauben, was man in pädagogischen Studientexten liest. (S/6) | 1 |

C: Item basiert auf Item von Conley et al. (2004, S. 202 f.) / Itemnummer

H: Item basiert auf Item von Hofer (2000, 390 f.) / Itemnummer

SDB: Item basiert auf Item von Schraw et al. (2002, S. 275) / Itemnummer

U: Item basiert auf Item von Urhahne & Hopf (2004, S. 84 f.) / Itemnummer

S: Item basiert auf Item von Schommer (1998, 561 f.) / Itemnummer

TABELLE B.11: Itemwortlaut, -polung und -ursprung der 98-Item StEB-Version – Teil 4

| Item-code | Itemwortlaut | Pol-ung |
|---|---|---|
| qw_29 | Wenn mir beim Lesen eines pädagogischen Studientextes auffällt, dass dieser meiner Erfahrung nicht entspricht, dann bin ich mir sicher, mich wohl geirrt zu haben. (H/2) | 1 |
| qw_30 | Wenn ich die in einem pädagogischen Studientext beschriebene Position kenne, bin ich mir sicher, mich in dem Forschungsbereich gut auszukennen. (H/3) | 1 |
| qw_31 | Manchmal muss man die in einem pädagogischen Studientext gegebene Antwort einfach akzeptieren, auch wenn man sie nicht versteht. (H/4) | 1 |
| rw_2 | Mir ist es sehr wichtig, dass sich die Inhalte von pädagogischen Studientexten für mich richtig anfühlen. | 1 |
| rw_3 | Es ist mir unwichtig, ob mir Inhalte von pädagogischen Studientexten sofort plausibel erscheinen. | 0 |
| rw_4 | Es ist mir wichtig, dass die in pädagogischen Studientexten dargestellten Meinungen durch ausgewiesene Autoritäten legitimiert sind. | 1 |
| rw_5 | Ich muss nachvollziehen können, wie genau eine pädagogische Erkenntnis im Studientext zustande kam, um ihr Glauben zu schenken. | 0 |
| rw_7 | Für mich ist es schon sehr wichtig, welche Person einen pädagogischen Studientext geschrieben hat, wenn ich dem Inhalt Glauben schenken soll. | 1 |
| rw_8 | Pädagogische Studientexte werden von meinen Hochschullehrenden korrekt ausgewählt, so dass ich mich auf den Inhalt verlassen kann. | 1 |
| rw_11 | Es ist als Studierende/r wichtig, neben pädagogischen Studientexten noch weiteren Quellen zurate zu ziehen, um verschiedene Forschungsmeinungen zu konsultieren. | 0 |
| rw_12 | Es ist als Studierende/r wichtig, die für mein Studium geeigneten pädagogischen Studientextquellen auszuwählen und zu bewerten. | 0 |
| rw_13 | Da ich mich nicht einfach auf einzelne pädagogische Studientexte verlassen kann, kommt mir bei der Konstruktion meines Wissens eine wesentliche Aufgabe zu. | 0 |
| rw_18 | Wenn ein pädagogischer Studientext aus einem auflagenstarken Lehrbuch kommt, dann kann man sich sicher darauf verlassen, dass der Inhalt vertrauenswürdig ist. | 1 |
| rw_19 | Ich verlasse mich darauf, dass mir die wichtigen Inhalte von pädagogischen Studientexten in der Regel spontan plausibel sind. | 1 |
| rw_22 | Gute Theorien in pädagogischen Studientexten erkennt man daran, dass sie sich auf viele verschiedene Forschungsergebnisse stützen. (C/22, U/24) | 0 |
| rw_23 | Um einschätzen zu können, ob die in einem pädagogischen Studientext dargestellte Meinung verlässlich ist, ist es wichtig, dass ich versuche Verbindungen zu meinem Vorwissen herzustellen. | 0 |

C: Item basiert auf Item von Conley et al. (2004, S. 202 f.) / Itemnummer

H: Item basiert auf Item von Hofer (2000, 390 f.) / Itemnummer

U: Item basiert auf Item von Urhahne & Hopf (2004, S. 84 f.) / Itemnummer

TABELLE B.12: Itemwortlaut, -polung und -ursprung der 98-Item StEB-Version – Teil 5

| Item-code | Itemwortlaut | Pol-ung |
|---|---|---|
| rw_24 | Wenn ich einen pädagogischen Studientext lese, ist es mir wichtig, mir ein eigenes Erklärungsmodell der Inhalte zu bauen. | 0 |
| rw_25 | Mir ist es wichtig einzuschätzen, ob das, was ich lese, in sich logisch ist, um beim Lesen eines pädagogischen Studientextes zu Erkenntnissen zu gelangen. | 0 |
| rw_26 | Den in einem pädagogischen Studientext aufgestellten Behauptungen kann man erst trauen, wenn man sich mit mehreren Wissensquellen zu dem Thema vertraut gemacht hat. | 0 |
| rw_27 | Es ist mir wichtig einen pädagogischen Studientext nicht nur zu lesen, sondern selbst zu durchdenken. | 0 |
| rw_28 | Ich bilde mir meine Meinung zu pädagogischen Themen nicht auf Basis eines einzelnen Studientextes. | 0 |
| rw_29 | Wenn die Autorin oder der Autor eines pädagogischen Studientextes eine hohe Position im Fachgebiet inne hat, dann bin ich mir sicher, dass zutrifft, was sie/er schreibt. | 1 |
| rw_30 | Ich vertraue den Inhalten von pädagogischen Studientexten, auch wenn die Autorinnen und Autoren keine Begründungen und Belege anführen. | 1 |
| rw_31 | Pädagogische Studientexte, in denen die Autorin oder ein Autor nur Behauptungen aufstellt, ohne diese zu belegen oder gut zu begründen, halte ich für weniger vertrauenswürdig. | 0 |
| rw_32 | Nicht pädagogische Studientexte, sondern eigene Erfahrung ist der beste Weg, um etwas in diesem Fach herauszufinden. (H/13) | 1 |
| rw_33 | Da ich nicht prüfen kann, ob stimmt was ich in pädagogischen Studientexten lese, glaube ich es erstmal. | 1 |
| rw_35* | Für die Vertrauenswürdigkeit des enthaltenen Wissens spielt es eine wichtige Rolle, wann ein pädagogischer Studientext publiziert wurde. | 0 |
| rw_36* | Gute pädagogische Studientexte sind diejenigen, mit denen man die Prüfungen einfach und gut besteht. | 1 |
| rw_37* | Es ist wichtig, dass Studierende über eigene Methoden der Überprüfung verfügen, um die Reichweite des Wissens in pädagogischen Studientexten einschätzen zu können. | 0 |
| sw_1 | Die Dinge sind einfacher als sie in den meisten pädagogischen Studientexten dargestellt werden. | 1 |
| sw_2 | Es ist wichtig, dass die Sachverhalte in pädagogischen Studientexten einfach dargestellt werden. | 1 |
| sw_3 | Ich bin der Überzeugung, dass Autorinnen und Autoren, die die Sachverhalte in pädagogischen Studientexten komplex darstellen, sich eigentlich gar nicht so gut auskennen. | 1 |
| sw_4 | Wenn ein Sachverhalt komplex ist, dann führt auch kein Weg daran vorbei, ihn auch in seiner ganzen Komplexität im pädagogischen Studientext darzustellen. | 0 |

* Veränderter Itemcode aufgrund von veränderter Zuordnung des Items zu anderer Subskala: rw_35 vor Expertenbefragung ursprünglich qw_16; rw_36 vor Expertenbefragung ursprünglich qw_17; rw_37 vor Expertenbefragung ursprünglich qw_7.

H: Item basiert auf Item von Hofer (2000, 390 f.) / Itemnummer

TABELLE B.13: Itemwortlaut, -polung und -ursprung der 98-Item StEB-Version – Teil 6

| Itemcode | Itemwortlaut | Polung |
|---|---|---|
| sw_7 | Es ist wichtig, dass pädagogische Studientexte die verschiedenen Sachverhalte streng nach einzelnen Inhalten getrennt darstellen. | 1 |
| sw_8 | Es ist wichtig, dass Überschneidungen von Inhaltsbereichen in pädagogischen Studientexten vermieden werden. | 1 |
| sw_9 | Es ist wichtig, dass pädagogische Studientexte viele Inhaltsbereiche verbinden und diese zusammenführen. | 0 |
| sw_10 | Es ist wichtig, dass pädagogische Studientexte auf die verschiedenartigen Verbindungen zu inhaltlichen Nachbarfeldern eingehen. | 0 |
| sw_11 | Ich bin der Überzeugung, dass sich wirklich nützliches Wissen in pädagogischen Studientexten in der Regel sehr einfach darstellen lässt. | 1 |
| sw_13 | Ich bin der Überzeugung, dass Wissen in pädagogischen Studientexten nützlich ist, wenn es wie ein Rezept organisiert ist: Es werden in der Reihenfolge der Handlungen alle wesentlichen Schritte genannt. | 1 |
| sw_14 | Das Wissen, das in pädagogischen Studientexten behandelt wird, ist komplex und sogar teilweise widersprüchlich, weil die echte Welt so vielfältig ist. | 0 |
| sw_15 | In pädagogischen Studientexten werden Dinge oft unnötig verkompliziert. | 1 |
| sw_16 | Die echte Welt ist viel komplexer als dies in unseren pädagogischen Studientexten oft dargestellt wird. | 0 |
| sw_17 | Es ist mir wichtig, dass Wissen aus pädagogischen Studientexten immer einfach anzuwenden ist. | 1 |
| sw_18 | Wirklich anwendbares Wissen muss in pädagogischen Studientexten eher komplex dargestellt werden. | 0 |
| sw_19 | Wirklich relevantes Wissen ist in pädagogischen Studientexten immer widerspruchsfrei. | 1 |
| sw_20 | Brauchbares Wissen wird in pädagogischen Studientexten immer in einfache Ursache-Wirkung-Beziehungen herunter gebrochen. | 1 |
| sw_21 | Wenn man sich die Fakten aus einem pädagogischen Studientext aneignet, dann weiß man das Wichtigste. | 1 |
| sw_22 | Es ist wichtig, die Informationen eines pädagogischen Studientextes mit denen aus anderen pädagogischen Studientexten zu verbinden. (S/14) | 0 |
| sw_23 | Es ist wichtig beim Lesen eines pädagogischen Studientextes nach den spezifischen Fakten zu suchen. (S/38) | 1 |
| sw_24 | Einen pädagogischen Studientext verstehen heißt, eher das große Ganze zu verstehen als die Details. (S/18) | 0 |
| sw_27* | Die meisten Autorinnen und Autoren von pädagogischen Studientexten stellen die Themen komplizierter dar, als sie eigentlich sind. (SBD/17) | 1 |

* Veränderter Itemcode aufgrund von veränderter Zuordnung des Items zu anderer Subskala: sw_27 vor Expertenbefragung ursprünglich gw_14.

SDB: Item basiert auf Item von Schraw et al. (2002, S. 275) / Itemnummer

S: Item basiert auf Item von Schommer (1998, 561 f.) / Itemnummer

## 34-Item StEB-Version

TABELLE B.14: Code, Wortlaut und Polung der Items der 34-Item StEB-Version – Teil 1

| Item-code | Itemwortlaut | Pol-ung |
|---|---|---|
| rw_33_u | Da ich nicht prüfen kann, ob stimmt was ich in pädagogischen Studientexten lese, glaube ich es erstmal. | 1 |
| qw_25_u | Ich zweifle nicht an, was in pädagogischen Studientexten steht. | 1 |
| qw_21_u | Was man in einem pädagogischen Studientext liest, ist in der Regel wahr. | 1 |
| qw_26_u | Ich stelle die Autorität der Autorin oder des Autors eines pädagogischen Studientextes nicht in Frage. | 1 |
| qw_13_u | Ich verlasse mich darauf, dass in einem pädagogischen Studientext steht, was in meinem Fach genau gilt und daran halte ich mich. | 1 |
| rw_8_u | Pädagogische Studientexte werden von meinen Hochschullehrenden korrekt ausgewählt, so dass ich mich auf den Inhalt verlassen kann. | 1 |
| gw_4_u | Ich verlasse mich auf Forschungserkenntnisse in pädagogischen Studientexten, wenn zugehörige Studien genannt werden. | 1 |
| rw_18_u | Wenn ein pädagogischer Studientext aus einem auflagenstarken Lehrbuch kommt, dann kann man sich sicher darauf verlassen, dass der Inhalt vertrauenswürdig ist. | 1 |
| rw_29_u | Wenn die Autorin oder der Autor eines pädagogischen Studientextes eine hohe Position im Fachgebiet inne hat, dann bin ich mir sicher, dass zutrifft, was sie/er schreibt. | 1 |
| sw_15_u | In pädagogischen Studientexten werden Dinge oft unnötig verkompliziert. | 1 |
| gw_13_u | Ein pädagogischer Studientext sollte die richtige Antwort, d.h. Fakten enthalten, statt viele Ansichten und Theorien zu diskutieren. | 1 |
| gw_3_u | Es ist Aufgabe der Expertinnen und Experten, die inhaltlichen Differenzen im Wissen auszudiskutieren. Erst danach sollten sie pädagogische Studientexte verfassen. | 1 |
| sw_2_u | Es ist wichtig, dass die Sachverhalte in pädagogischen Studientexten einfach dargestellt werden. | 1 |
| sw_3_u | Ich bin der Überzeugung, dass Autorinnen und Autoren, die die Sachverhalte in pädagogischen Studientexten komplex darstellen, sich eigentlich gar nicht so gut auskennen. | 1 |
| sw_11_u | Ich bin der Überzeugung, dass sich wirklich nützliches Wissen in pädagogischen Studientexten in der Regel sehr einfach darstellen lässt. | 1 |
| sw_17_u | Es ist mir wichtig, dass Wissen aus pädagogischen Studientexten immer einfach anzuwenden ist. | 1 |
| rw_36_u | Gute pädagogische Studientexte sind diejenigen, mit denen man die Prüfungen einfach und gut besteht. | 1 |

TABELLE B.15: Code, Wortlaut und Polung der Items der 34-Item StEB-Version – Teil 2

| Item-code | Itemwortlaut | Pol-ung |
|---|---|---|
| qw_11 | Für mein eigenes Wissen ist es wichtig, mir über möglichst viele verschiedene Ansichten aus verschiedenen pädagogischen Studientexten einen Überblick zu verschaffen. | 0 |
| gw_9 | Es ist wichtig, dass pädagogische Studientexte neben Herleitungen und Aussagen auch Auskunft darüber geben, wie die einzelnen Aspekte genau erforscht wurden. | 0 |
| sw_9 | Es ist wichtig, dass pädagogische Studientexte viele Inhaltsbereiche verbinden und diese zusammenführen. | 0 |
| rw_27 | Es ist mir wichtig einen pädagogischen Studientext nicht nur zu lesen, sondern selbst zu durchdenken. | 0 |
| gw_16 | Es kommt bei einem pädagogischen Studientext darauf an, dass viele unterschiedliche Ansichten diskutiert werden. | 0 |
| rw_12 | Es ist als Studierende/r wichtig, die für mein Studium geeigneten pädagogischen Studientextquellen auszuwählen und zu bewerten. | 0 |
| rw_11 | Es ist als Studierende/r wichtig, neben pädagogischen Studientexten noch weiteren Quellen zurate zu ziehen, um verschiedene Forschungsmeinungen zu konsultieren. | 0 |
| sw_22 | Es ist wichtig, die Informationen eines pädagogischen Studientextes mit denen aus anderen pädagogischen Studientexten zu verbinden. | 0 |
| gw_17 | Ich bin offen gegenüber Ansichten in pädagogischen Studientexten, die ich bisher noch nicht kannte. | 0 |
| ew_10 | Forschende entdecken sehr oft Relevantes und Neues, so dass ich mein pädagogisches Wissen in kürzerer Zeit mittels neuer Studientexte immer wieder aktualisieren sollte. | 0 |
| ew_13 | Die Vorstellungen, die in pädagogischen Studientexten dargestellt werden, verändern sich über die Zeit. | 0 |
| ew_12 | Viele Ansichten in heutigen pädagogischen Studientexten entsprechen nicht mehr dem, was früher als richtig angenommen wurde. | 0 |
| ew_2 | Weil sich Wissen immer stark verändern wird, ist es völlig normal, wenn sich verschiedene pädagogische Studientexte inhaltlich widersprechen. | 0 |
| ew_16_u | Was heute in pädagogischen Studientexten als wahr gilt, trifft in Zukunft auch noch zu. | 1 |
| ew_5_u | Bei pädagogischen Studientexten kommen über die Jahre zwar ein paar Dinge hinzu aber dennoch bleiben die Erkenntnisse weitestgehend erhalten. | 1 |
| ew_3 | Wie etwas verstanden wird, kann sich in der pädagogischen Forschung sehr schnell ändern. Demnach ändert sich dann auch, was in Studientexten berichtet wird. | 0 |
| ew_8 | Ich werde mich schon in wenigen Jahren den meisten Inhalten aus den pädagogischen Studientexten neu widmen müssen, weil viele Dinge wieder überholt sein werden. | 0 |

## Finales 26-Item StEB Inventar

TABELLE B.16: Code, Wortlaut und Skalenzugehörigkeit der Items des finalen StEB Inventars – Teil 1

| neuer Code | alter Itemcode | Skala | Itemwortlaut |
|---|---|---|---|
| A1 | rw_33_u | Absolutheit | Da ich nicht prüfen kann, ob stimmt was ich in pädagogischen Studientexten lese, glaube ich es erstmal. |
| A2 | rw_29_u | Absolutheit | Wenn die Autorin oder der Autor eines pädagogischen Studientextes eine hohe Position im Fachgebiet inne hat, dann bin ich mir sicher, dass zutrifft, was sie/er schreibt. |
| A3 | qw_25_u | Absolutheit | Ich zweifle nicht an, was in pädagogischen Studientexten steht. |
| A4 | qw_21_u | Absolutheit | Was man in einem pädagogischen Studientext liest, ist in der Regel wahr. |
| A5 | qw_26_u | Absolutheit | Ich stelle die Autorität der Autorin oder des Autors eines pädagogischen Studientextes nicht in Frage. |
| A6 | qw_13_u | Absolutheit | Ich verlasse mich darauf, dass in einem pädagogischen Studientext steht, was in meinem Fach genau gilt und daran halte ich mich. |
| A7 | gw_4_u | Absolutheit | Ich verlasse mich auf Forschungserkenntnisse in pädagogischen Studientexten, wenn zugehörige Studien genannt werden |
| A8 | rw_18_u | Absolutheit | Wenn ein pädagogischer Studientext aus einem auflagenstarken Lehrbuch kommt, dann kann man sich sicher darauf verlassen, dass der Inhalt vertrauenswürdig ist. |
| S1 | sw_15_u | Einfachheit | In pädagogischen Studientexten werden Dinge oft unnötig verkompliziert. |
| S2 | gw_13_u | Einfachheit | Ein pädagogischer Studientext sollte die richtige Antwort, d.h. Fakten enthalten, statt viele Ansichten und Theorien zu diskutieren. |
| S3 | sw_2_u | Einfachheit | Es ist wichtig, dass die Sachverhalte in pädagogischen Studientexten einfach dargestellt werden. |
| S4 | sw_17_u | Einfachheit | Es ist mir wichtig, dass Wissen aus pädagogischen Studientexten immer einfach anzuwenden ist. |
| S5 | rw_36_u | Einfachheit | Gute pädagogische Studientexte sind diejenigen, mit denen man die Prüfungen einfach und gut besteht. |
| S6 | gw_3_u | Einfachheit | Es ist Aufgabe der Expertinnen und Experten, die inhaltlichen Differenzen im Wissen auszudiskutieren. Erst danach sollten sie pädagogische Studientexte verfassen |

TABELLE B.17: Code, Wortlaut und Skalenzugehörigkeit der Items des finalen StEB Inventars – Teil 2

| neuer Code | Itemcode | Skala | Itemwortlaut |
|---|---|---|---|
| M1 | qw_11 | Multimodalität | Für mein eigenes Wissen ist es wichtig, mir über möglichst viele verschiedene Ansichten aus verschiedenen pädagogischen Studientexten einen Überblick zu verschaffen. |
| M2 | gw_9 | Multimodalität | Es ist wichtig, dass pädagogische Studientexte neben Herleitungen und Aussagen auch Auskunft darüber geben, wie die einzelnen Aspekte genau erforscht wurden. |
| M3 | rw_27 | Multimodalität | Es ist mir wichtig einen pädagogischen Studientext nicht nur zu lesen, sondern selbst zu durchdenken. |
| M4 | rw_12 | Multimodalität | Es ist als Studierende/r wichtig, die für mein Studium geeigneten pädagogischen Studientextquellen auszuwählen und zu bewerten. |
| M5 | rw_11 | Multimodalität | Es ist als Studierende/r wichtig, neben pädagogischen Studientexten noch weiteren Quellen zurate zu ziehen, um verschiedene Forschungsmeinungen zu konsultieren. |
| M6 | sw_22 | Multimodalität | Es ist wichtig, die Informationen eines pädagogischen Studientextes mit denen aus anderen pädagogischen Studientexten zu verbinden. |
| M7 | gw_16 | Multimodalität | Es kommt bei einem pädagogischen Studientext darauf an, dass viele unterschiedliche Ansichten diskutiert werden. |
| D1 | ew_13 | Entwicklung | Die Vorstellungen, die in pädagogischen Studientexten dargestellt werden, verändern sich über die Zeit. |
| D2 | ew _12 | Entwicklung | Viele Ansichten in heutigen pädagogischen Studientexten entsprechen nicht mehr dem, was früher als richtig angenommen wurde. |
| D3 | ew _3 | Entwicklung | Wie etwas verstanden wird, kann sich in der pädagogischen Forschung sehr schnell ändern. Demnach ändert sich dann auch, was in Studientexten berichtet wird. |
| D4 | ew _8 | Entwicklung | Ich werde mich schon in wenigen Jahren den meisten Inhalten aus den pädagogischen Studientexten neu widmen müssen, weil viele Dinge wieder überholt sein werden. |
| D5 | ew_10 | Entwicklung | Forschende entdecken sehr oft Relevantes und Neues, so dass ich mein pädagogisches Wissen in kürzerer Zeit mittels neuer Studientexte immer wieder aktualisieren sollte. |

## Abgeleitete Items der StEB-Versionen mit Itemquelle

TABELLE B.18: Auf bestehenden Fragebögen basierende Items der 34-Item StEB-Version

| StEB-Itemcode/ Itemursprung | Itemwortlaut |
|---|---|
| qw_21 | Was man in einem pädagogischen Studientext liest, ist in der Regel wahr. |
| H/1 | If you read something in a textbook for this subject, you can be sure it's true. |
| C/4 | If you read something in a science book, you can be sure it's true. |
| U/4 | Was man in einem Naturwissenschaftsbuch liest, ist sicher wahr. |
| qw_26 | Ich stelle die Autorität der Autorin oder des Autors eines pädagogischen Studientextes nicht in Frage. |
| SBD/26 | People shouldn't question authority. |
| ew_12 | Viele Ansichten in heutigen pädagogischen Studientexten entsprechen nicht mehr dem, was früher als richtig angenommen wurde. |
| C/10 | Some ideas in science today are different than what scientists used to think. |
| U/12 | Einige Vorstellungen in den Naturwissenschaften sind heute anders als das, was Naturwissenschaftler früher dachten. |
| ew_13 | Die Vorstellungen, die in pädagogischen Studientexten dargestellt werden, verändern sich über die Zeit. |
| C/11 | The ideas in science books sometimes change. |
| U/13 | Die Vorstellungen in Naturwissenschaftsbüchern verändern sich manchmal. |
| ew_16 | Was heute in pädagogischen Studientexten als wahr gilt, trifft in Zukunft auch noch zu. |
| SDB/23 | What is true today will be true tomorrow. |
| H/12 | Truth is unchanging in this subject. |
| sw_22 | Es ist wichtig, die Informationen eines pädagogischen Studientextes mit denen aus anderen pädagogischen Studientexten zu verbinden. |
| S/14 | I try my best to combine information across chapters or even across classes. (INTEG -) |

C: Item basiert auf Item von Conley et al. (2004, S. 202 f.) / Itemnummer
H: Item basiert auf Item von Hofer (2000, 390 f.) / Itemnummer
SDB: Item basiert auf Item von Schraw et al. (2002, S. 275) / Itemnummer
U: Item basiert auf Item von Urhahne & Hopf (2004, S. 84 f.) / Itemnummer
S: Item basiert auf Item von Schommer (1998, 561 f.) / Itemnummer

TABELLE B.19: Auf bestehenden Fragebögen basierende Items der finalen 26-Item StEB-Version

| StEB-Itemcode/ Itemursprung | Itemwortlaut |
|---|---|
| qw_21 | Was man in einem pädagogischen Studientext liest, ist in der Regel wahr. |
| H/1 | If you read something in a textbook for this subject, you can be sure it's true. |
| C/4 | If you read something in a science book, you can be sure it's true. |
| U/4 | Was man in einem Naturwissenschaftsbuch liest, ist sicher wahr. |
| qw_26 | Ich stelle die Autorität der Autorin oder des Autors eines pädagogischen Studientextes nicht in Frage. |
| SBD/26 | People shouldn't question authority. |
| ew_12 | Viele Ansichten in heutigen pädagogischen Studientexten entsprechen nicht mehr dem, was früher als richtig angenommen wurde. |
| C/10 | Some ideas in science today are different than what scientists used to think. |
| U/12 | Einige Vorstellungen in den Naturwissenschaften sind heute anders als das, was Naturwissenschaftler früher dachten. |
| ew_13 | Die Vorstellungen, die in pädagogischen Studientexten dargestellt werden, verändern sich über die Zeit. |
| C/11 | The ideas in science books sometimes change. |
| U/13 | Die Vorstellungen in Naturwissenschaftsbüchern verändern sich manchmal. |
| sw_22 | Es ist wichtig, die Informationen eines pädagogischen Studientextes mit denen aus anderen pädagogischen Studientexten zu verbinden. |
| S/14 | I try my best to combine information across chapters or even across classes. (INTEG -) |

C: Item basiert auf Item von Conley et al. (2004, S. 202 f.) / Itemnummer
H: Item basiert auf Item von Hofer (2000, 390 f.) / Itemnummer
SDB: Item basiert auf Item von Schraw et al. (2002, S. 275) / Itemnummer
U: Item basiert auf Item von Urhahne & Hopf (2004, S. 84 f.) / Itemnummer
S: Item basiert auf Item von Schommer (1998, 561 f.) / Itemnummer

# C Lernstrategien im Studium (LIST-R)

TABELLE C.1: Itemwortlaut Lernstrategien im Studium (LIST-R, Wild (2011)) - Teil 1

| Item | Itemwortlaut |
|---|---|
| AL.IO1 | Ich fertige Tabellen, Diagramme oder Schaubilder an, um den Stoff besser strukturiert vorliegen zu haben. |
| AL.IO2 | Ich mache mir kurze Zusammenfassungen der wichtigsten Inhalte. |
| AL.IO3 | Ich mache mir eine Gliederung mit den wichtigsten Punkten. |
| AL.IO4 | Ich versuche, den Stoff systematisch zu ordnen. |
| AL.IO5 | Ich stelle mir aus dem Lernstoff kurze Zusammenfassungen mit den Hauptideen zusammen. |
| AL.IO6 | Für größere Stoffmengen fertige ich eine Gliederung an, die die Struktur des Stoffs am besten wiedergibt. |
| AL.IO7 | Ich stelle wichtige Fachausdrücke und Definitionen in eigenen Listen zusammen. |
| AL.IWE1 | Ich präge mir den Lernstoff durch Wiederholen ein. |
| AL.IWE2 | Ich lese meine Aufzeichnungen mehrmals hintereinander durch. |
| AL.IWE3 | Ich lerne Schlüsselbegriffe auswendig, um mich in der Prüfung besser an wichtige Inhaltsbereiche erinnern zu können. |
| AL.IWE4 | Ich lese einen Abschnitt und versuche dann, ihn auswendig wiederzugeben. |
| AL.IWE5 | Ich lerne Regeln, Fachbegriffe oder Formeln auswendig. |
| AL.IWE6 | Ich gehe in Gedanken noch einmal den gelernten Stoff durch, um zu sehen, ob ich alles Wesentliche behalten habe. |
| AL.IWE7 | Ich lerne den Lernstoff anhand von Skripten oder anderen Aufzeichnungen möglichst auswendig. |
| AL.IVP1 | Zu neuen Konzepten stelle ich mir praktische Anwendungen vor. |
| AL.IVP2 | Ich versuche, neue Begriffe oder Theorien auf mir bereits bekannte Begriffe und Theorien zu beziehen. |
| AL.IVP3 | Ich versuche in Gedanken, das Gelernte mit dem zu verbinden, was ich schon darüber weiß. |
| AL.IVP4 | Ich denke mir konkrete Beispiele zu bestimmten Lerninhalten aus. |
| AL.IVP5 | Ich prüfe, ob Theorien, Interpretationen oder Schlussfolgerungen ausreichend belegt und begründet sind. |
| AL.IVP6 | Ich denke über Alternativen zu den Behauptungen oder Schlussfolgerungen nach. |
| AL.IVP7 | Der neue Stoff dient mir als Ausgangspunkt für die Entwicklung eigener Ideen. |
| LO.MP1 | Ich lege im vornhinein fest, wie weit ich mit der Durcharbeitung des Stoffs kommen möchte. |
| LO.MP2 | Vor dem Lernen überlege ich mir, wie ich am effektivsten vorgehen kann. |
| LO.MP3 | Ich überlege mir, in welcher Reihenfolge ich den Stoff am besten durcharbeite. |

TABELLE C.2: Itemwortlaut Lernstrategien im Studium (LIST-R, Wild (2011)) - Teil 2

| Item | Itemwortlaut |
|---|---|
| LO.MQ1 | Ich stelle mir Fragen zum Stoff, um sicher zu gehen, dass ich alles verstanden habe. |
| LO.MQ2 | Um Wissenslücken festzustellen, rekapituliere ich die wichtigsten Inhalte, ohne meine Unterlagen zu Hilfe zu nehmen. |
| LO.MQ3 | Ich bearbeite zusätzliche Aufgaben, um festzustellen, ob ich den Stoff wirklich verstanden habe. |
| LO.MQ4 | Wenn ich Schwierigkeiten beim Lernen bestimmter Inhalte habe, gehe ich meine Arbeit noch einmal langsam durch. |
| LO.MQ5 | Wenn ich während des Lesens nicht alles verstehe, versuche ich die Lücken festzuhalten und den Text daraufhin noch einmal durchzugehen. |
| LO.MQ6 | Wenn mir eine bestimmte Textstelle unklar erscheint, versuche ich durch genaues Lesen die Unklarheit zu beseitigen. |
| LO.MQ7 | Wenn ich einen Sachverhalt beim ersten Mal nicht verstehe, gehe ich ihn noch einmal Schritt für Schritt durch. |
| RN.LZ1 | Beim Lernen halte ich mich an einen bestimmten Zeitplan. |
| RN.LZ2 | Ich lege bestimmte Zeiten fest, zu denen ich dann lerne. |
| RN.LZ3 | Ich lege die Stunden, die ich täglich mit Lernen verbringe, durch einen Zeitplan fest. |
| RN.LZ4 | Ich lege vor jeder Lernphase eine bestimmte Zeitdauer fest. |
| RN.LU1 | Ich lerne an einem Platz, an dem ich mich wohlfühle. |
| RN.LU2 | Ich lerne an einem Platz, an dem ich konzentriert arbeiten kann. |
| RN.LU3 | Ich gestalte meine Umgebung so, dass ich möglichst nicht abgelenkt werde. |
| RN.LU4 | Wenn ich lerne, sorge ich dafür, dass ich in Ruhe arbeiten kann. |
| RN.LU5 | Mein Arbeitsplatz ist so gestaltet, dass ich alles schnell finden kann. |
| RN.LU6 | Ich habe an meinem Arbeitsplatz die wichtigsten Unterlagen griffbereit. |
| RN.LN1 | Wenn mir ein Inhalt noch nicht ganz klar ist, suche ich nach ergänzender Literatur. |
| RN.LN2 | Wenn ich einen Begriff nicht verstehe, schlage ich in der Literatur nach. |
| RN.LN3 | Fehlende Informationen suche ich mir aus verschiedenen Quellen zusammen (z.B. Bücher, Fachzeitschriften). |
| RN.LN4 | Bei Bedarf ziehe ich zusätzliche Literatur heran. |
| RN.OQ1 | Wenn ich zusätzliche Literatur suche, benutze ich entsprechende Datenbanken im Netz. |
| RN.OQ2 | Ich nutze für mein Studium Online Lexika (z.B. Wikipedia). |
| RN.OQ3 | Ich benutze bei Bedarf Suchmaschinen (Google, Yahoo, etc.). |
| RN.OQ4 | Ich nutze Infos aus speziellen Internet-Foren. |
| RN.GL1 | Ich lerne gemeinsam mit Studienkolleg(inn)en. |
| RN.GL2 | Ich nehme mir gerne Zeit, um mit Studienkolleg(inn)en zu diskutieren. |
| RN.GL3 | Wenn ich Verständnisprobleme habe, nehme ich die Hilfe von Studienkolleg(inn)en in Anspruch. |
| RN.GL4 | Wenn mir etwas unklar ist, wende ich mich an meine Studienkolleg(inn)en. |

# D CFA-Modelle zur Studie 2

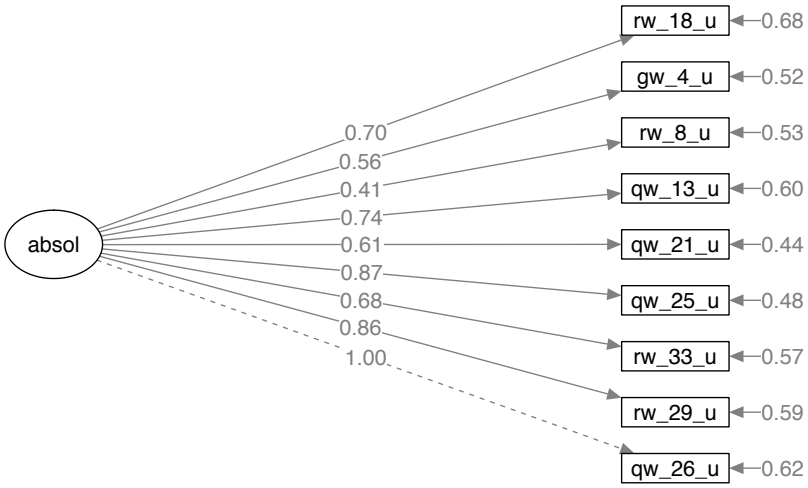


ABBILDUNG D.1: Initiales, unstandardisiertes, einfaktorielles Modell des Faktors Absolutheit ($N$ = 915)

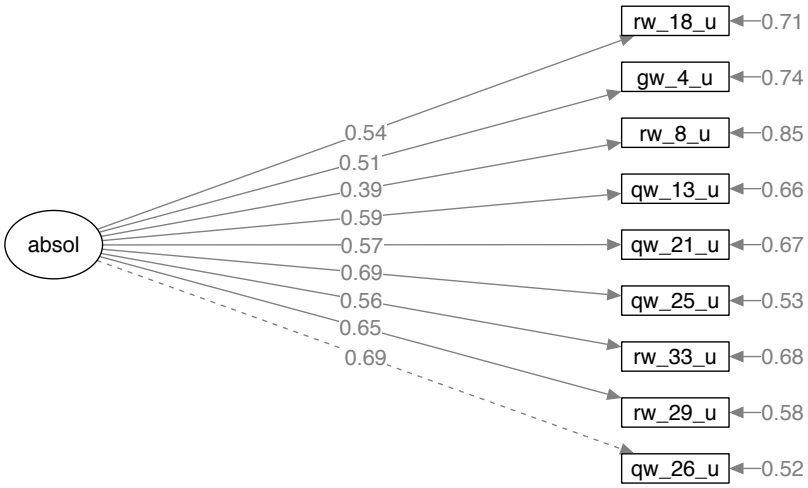


ABBILDUNG D.2: Initiales, standardisiertes, einfaktorielles Modell des Faktors Absolutheit ($N$ = 915)

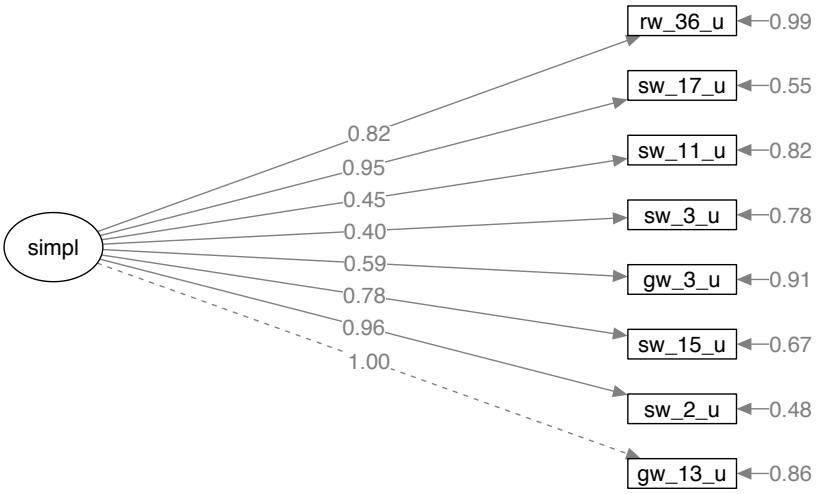


ABBILDUNG D.3: Initiales, unstandardisiertes, einfaktorielles Modell des Faktors Einfachheit ($N$ = 915)

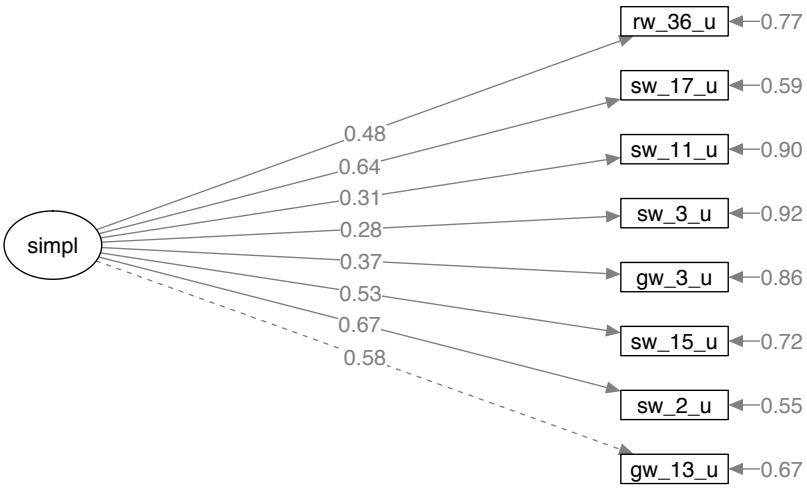


ABBILDUNG D.4: Initiales, standardisiertes, einfaktorielles Modell des Faktors Einfachheit ($N$ = 915)

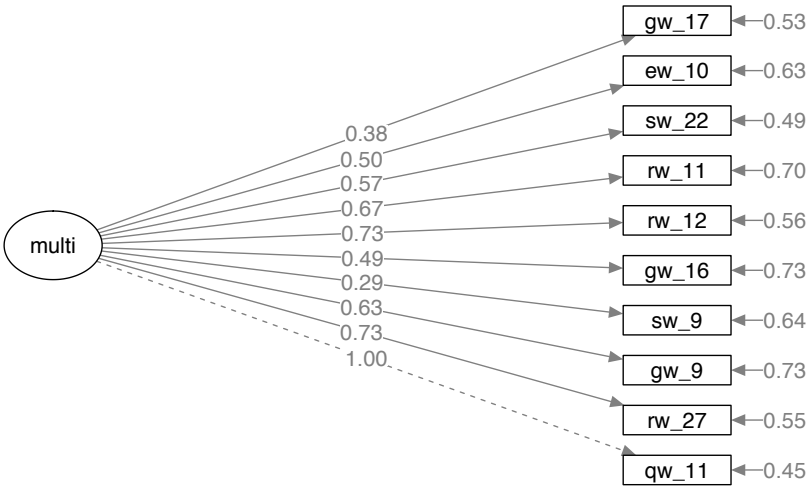


ABBILDUNG D.5: Initiales, unstandardisiertes, einfaktorielles Modell des Faktors Multimodalität ($N$ = 915)

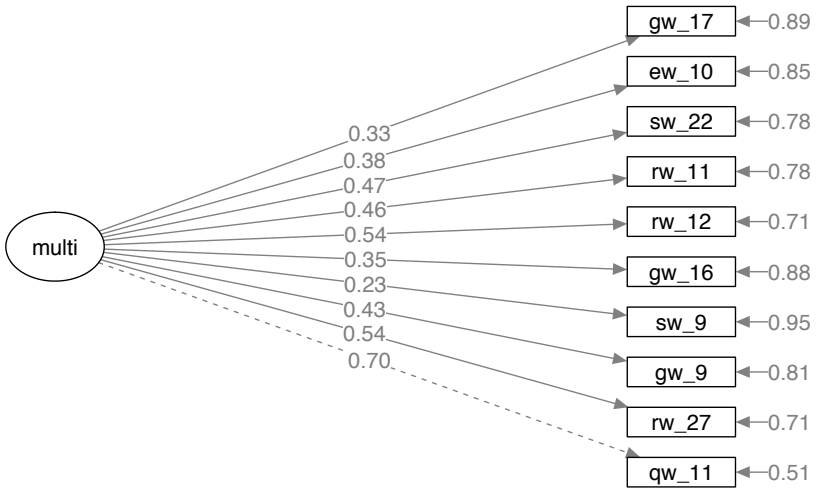


ABBILDUNG D.6: Initiales, standardisiertes, einfaktorielles Modell des Faktors Multimodalität ($N$ = 915)

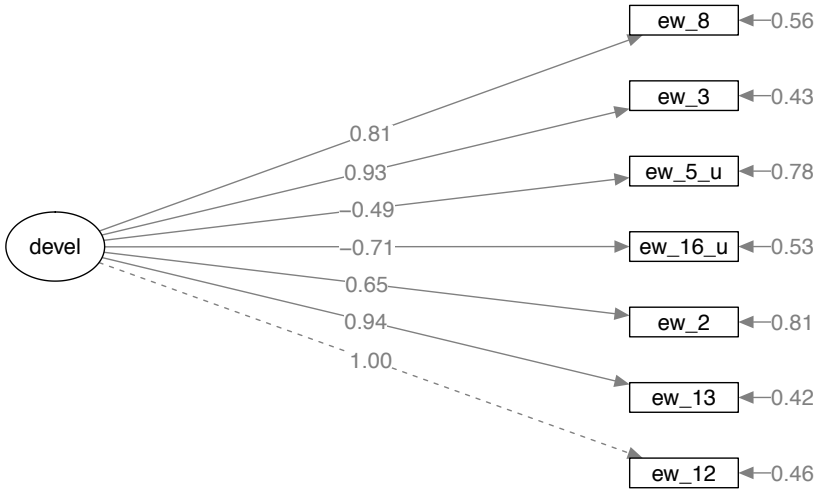


ABBILDUNG D.7: Initiales, unstandardisiertes, einfaktorielles Modell des Faktors Entwicklung ($N$ = 915)

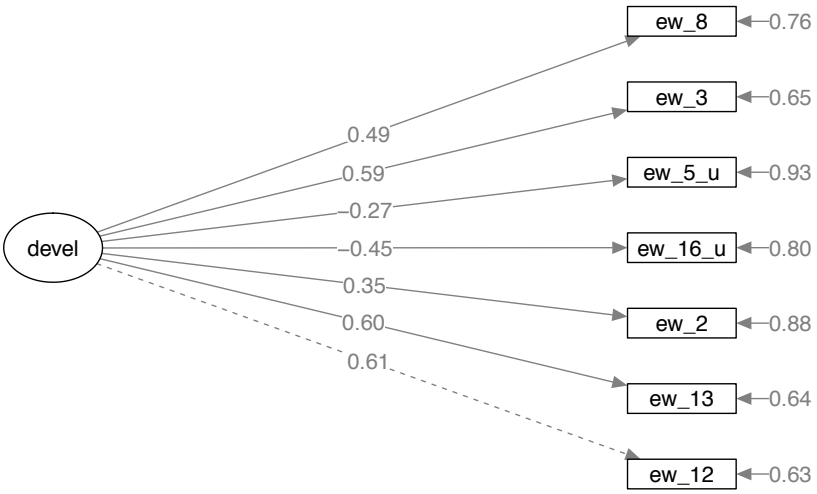


ABBILDUNG D.8: Initiales, standardisiertes, einfaktorielles Modell des Faktors Entwicklung ($N$ = 915)

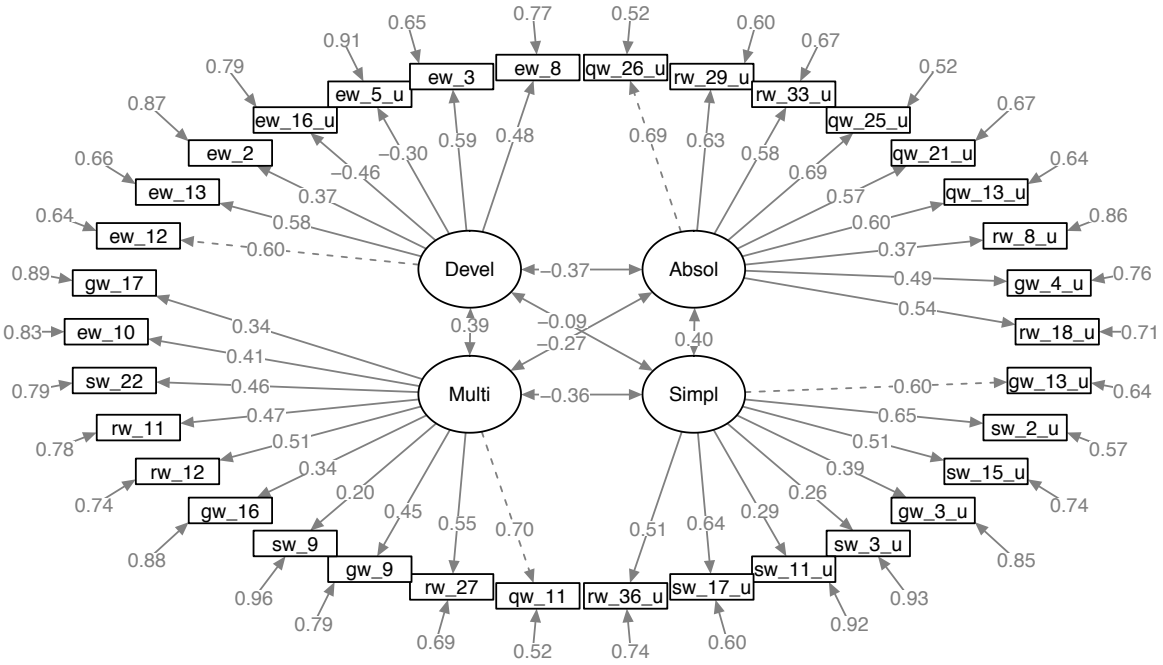


ABBILDUNG D.9: Standardisiertes vierfaktorielles Modell ($N$ = 915)

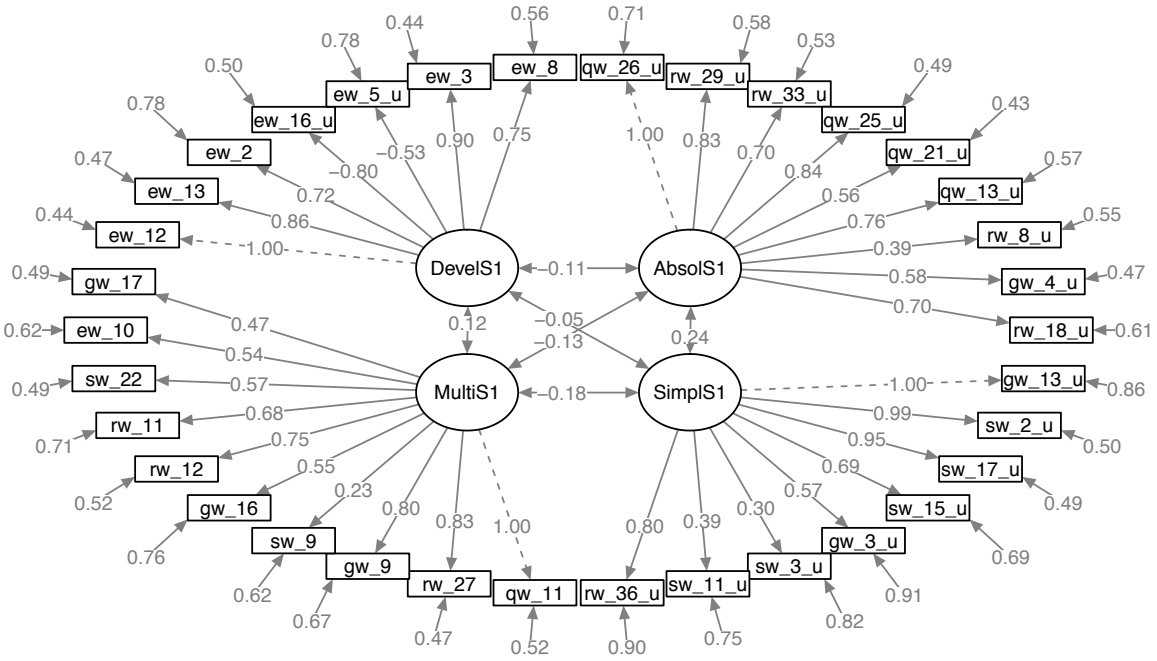


ABBILDUNG D.10: Initiales, unstandardisiertes vierfaktorielles Modell zu Teilstichprobe 1 ($N$ = 457)

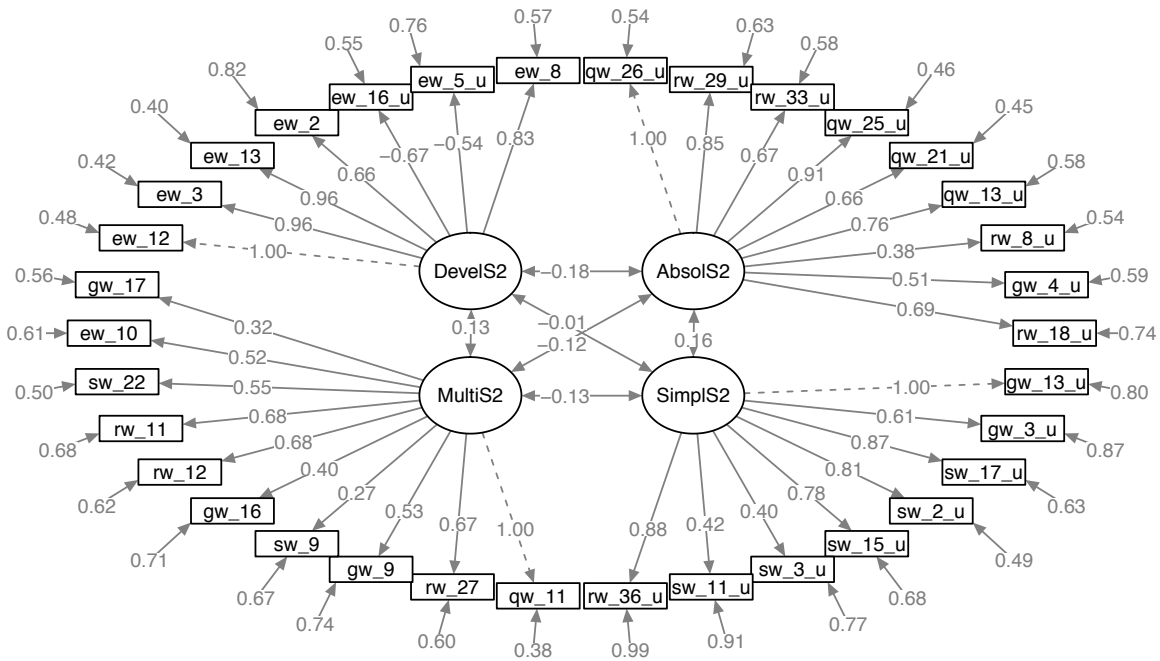


ABBILDUNG D.11: Initiales, unstandardisiertes vierfaktorielles Modell zu Teilstichprobe 2 ($N$ = 458)

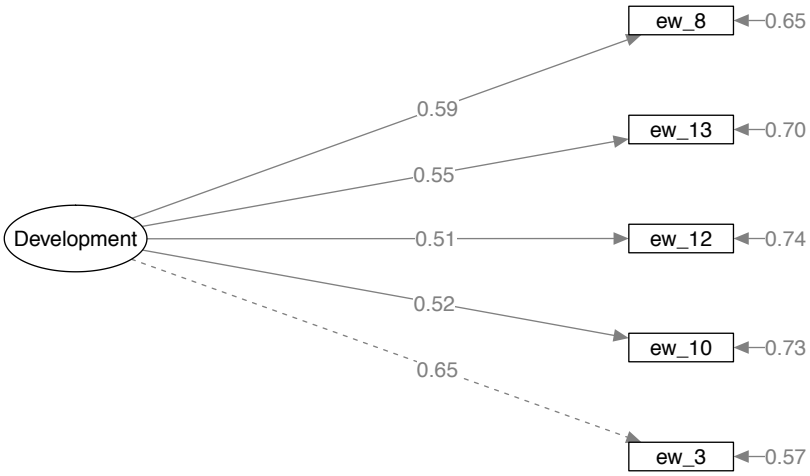


ABBILDUNG D.12: Finales, standardisiertes einfaktorielles Modell des Faktors Entwicklung ($N$ = 915)

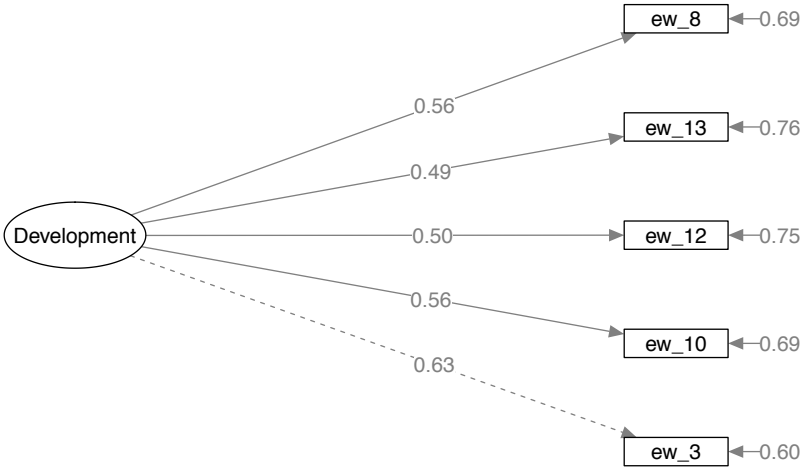


ABBILDUNG D.13: Finales, standardisiertes einfaktorielles Modell des Faktors Entwicklung für Teilstichprobe 1 ($N$ = 457)

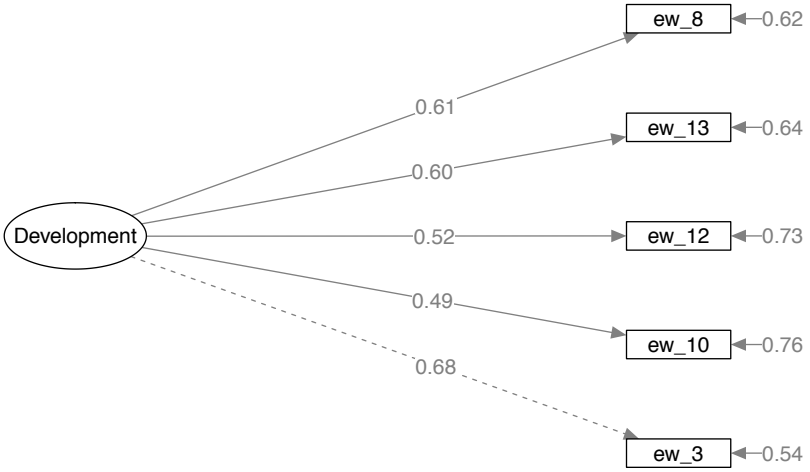


ABBILDUNG D.14: Finales, standardisiertes einfaktorielles Modell des Faktors Entwicklung für Teilstichprobe 2 ($N$ = 458)

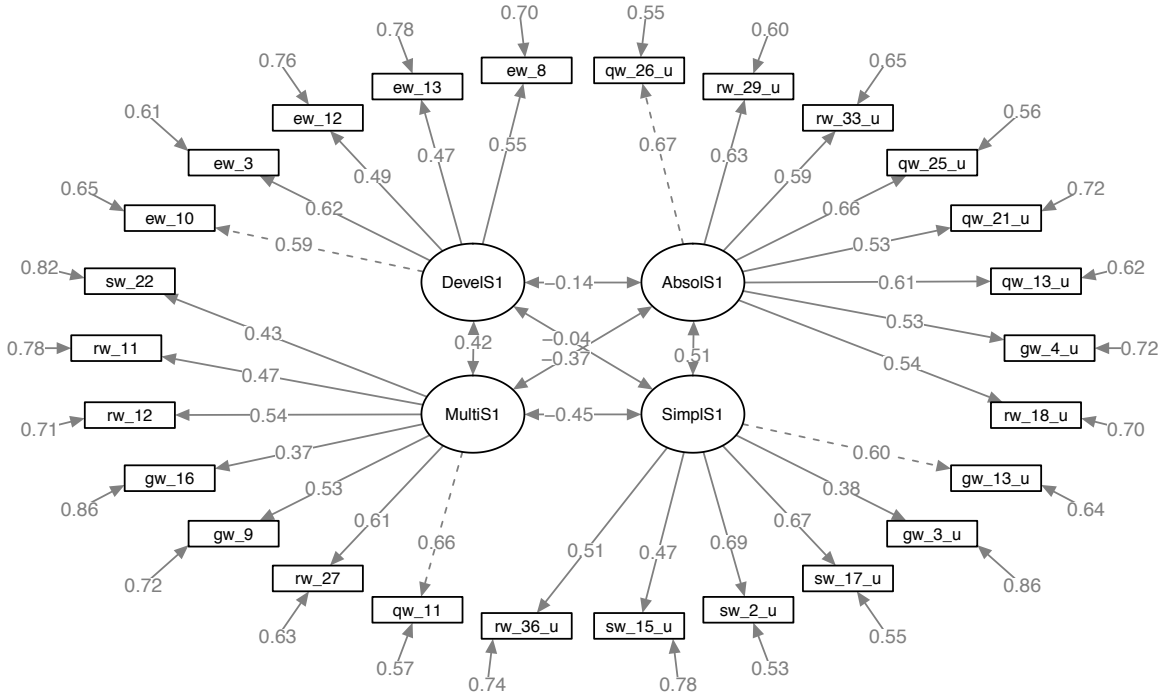


ABBILDUNG D.15: Finales, standardisiertes 26-Item Modell Teilstichprobe 1 ($N$ = 457)

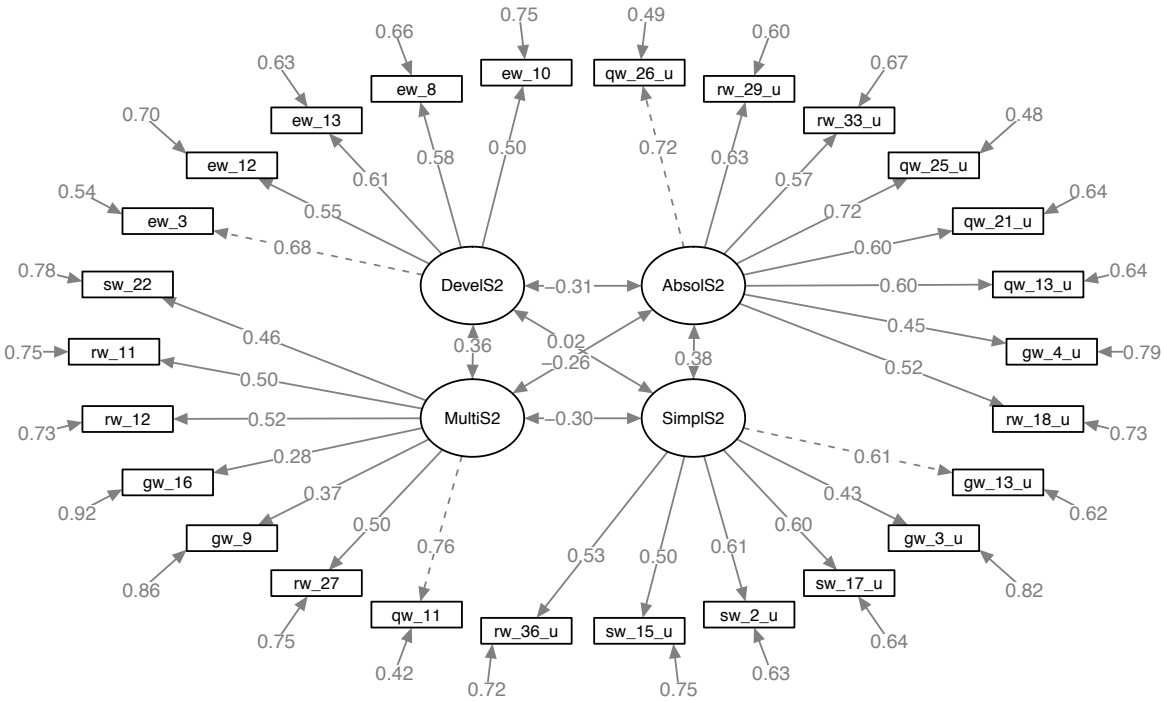


ABBILDUNG D.16: Finales, standardisiertes 26-Item Modell Teilstichprobe 2 ($N$ = 458)

# E Parameterschätzungen zur Studie 2

## Parameterschätzungen des finalen 26-Item vierfaktoriellen Gesamtmodells mit korrelierten Fehlervarianzen

TABELLE E.1: Geschätzte semipartielle Regressionsgewichte (Ladungen $\lambda_{i1}$) der Items des finalen 26-Item vierfaktoriellen Modells mit korrelierten Itemfehlervarianzen ($N$ = 915)

| Factor regressed on item | geschätzte $\lambda$ | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| Absolutheit mit | | | | | |
| qw_26_u | 1.000 | | | | .694 |
| rw_29_u | .789 | .049 | 16.206 | .000 | .599 |
| rw_33_u | .712 | .047 | 15.162 | .000 | .592 |
| qw_25_u | .885 | .044 | 20.155 | .000 | .702 |
| qw_21_u | .602 | .045 | 13.255 | .000 | .564 |
| qw_13_u | .765 | .045 | 16.818 | .000 | .609 |
| gw_4_u | .536 | .043 | 12.393 | .000 | .485 |
| rw_18_u | .642 | .051 | 12.649 | .000 | .499 |
| Einfachheit mit | | | | | |
| gw_13_u | 1.000 | | | | .608 |
| gw_3_u | .600 | .063 | 9.522 | .000 | .404 |
| sw_17_u | .871 | .070 | 12.430 | .000 | .622 |
| sw_2_u | .850 | .068 | 12.501 | .000 | .630 |
| sw_15_u | .676 | .063 | 10.740 | .000 | .484 |
| rw_36_u | .853 | .072 | 11.925 | .000 | .523 |
| Multimodalität mit | | | | | |
| qw_11 | 1.000 | | | | .696 |
| rw_27 | .747 | .059 | 12.743 | .000 | .553 |
| gw_9 | .671 | .067 | 10.007 | .000 | .462 |
| gw_16 | .466 | .057 | 8.154 | .000 | .333 |
| rw_12 | .708 | .052 | 13.527 | .000 | .519 |
| rw_11 | .694 | .059 | 11.795 | .000 | .480 |
| sw_22 | .538 | .057 | 9.432 | .000 | .444 |
| Entwicklung mit | | | | | |
| ew_3 | 1.000 | | | | .689 |
| ew_10 | .765 | .084 | 9.122 | .000 | .497 |
| ew_12 | .735 | .076 | 9.625 | .000 | .484 |
| ew_13 | .749 | .074 | 10.065 | .000 | .517 |
| ew_8 | .833 | .068 | 12.243 | .000 | .543 |

TABELLE E.2: Geschätzte Kovarianzen zwischen den Items und Faktoren des finalen 26-Item vierfaktoriellen Modells mit korrelierten Itemfehlervarianzen ($N$ = 915)

| Faktoren | Geschätzte Kovarianz | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| rw_29_u mit rw_18_u | .129 | .030 | 4.353 | .000 | .190 |
| sw_17_u mit sw_22 | .097 | .023 | 4.223 | .000 | .181 |
| ew_12 mit ew_13 | .083 | .024 | 3.511 | .000 | .161 |
| ew_10 mit ew_8 | .086 | .025 | 3.422 | .001 | .161 |
| rw_29_u mit sw_2_u | -.097 | .022 | -4.321 | .000 | -.168 |
| Absolutheit mit | | | | | |
| Einfachheit | .247 | .031 | 8.073 | .000 | .472 |
| Multimodalität | -.161 | .025 | -6.358 | .000 | -.326 |
| Entwicklung | -.105 | .023 | -4.551 | .000 | -.248 |
| Einfachheit mit | | | | | |
| Multimodalität | -.183 | .027 | -6.814 | .000 | -.406 |
| Entwicklung | -.000 | .019 | -0.010 | 0.992 | -.000 |
| Multimodalität mit | | | | | |
| Entwicklung | .141 | .019 | 7.455 | .000 | .386 |

TABELLE E.3: Geschätzte $R^2$ der Items des finalen 26-Item vierfaktoriellen Modells mit korrelierten Itemfehlervarianzen ($N$ = 915)

| Items | $R^2$ | Items | $R^2$ |
|---|---|---|---|
| qw_26_u | .481 | rw_36_u | .274 |
| rw_29_u | .359 | qw_11 | .485 |
| rw_33_u | .351 | rw_27 | .306 |
| qw_25_u | .493 | gw_9 | .214 |
| qw_21_u | .318 | gw_16 | .111 |
| qw_13_u | .371 | rw_12 | .270 |
| gw_4_u | .235 | rw_11 | .230 |
| rw_18_u | .249 | sw_22 | .197 |
| gw_13_u | .370 | ew_3 | .474 |
| gw_3_u | .163 | ew_10 | .247 |
| sw_17_u | .387 | ew_12 | .234 |
| sw_2_u | .397 | ew_13 | .267 |
| sw_15_u | .234 | ew_8 | .295 |

TABELLE E.4: Geschätzte Intercepts (Itemmittelwerte) der Items des finalen 26-Item vierfaktoriellen Modells mit korrelierten Itemfehlervarianzen ($N$ = 915)

| Items | Geschätzte Intercepts | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| qw_26_u | 2.957 | .036 | 81.762 | .000 | 2.703 |
| rw_29_u | 3.077 | .033 | 92.519 | .000 | 3.079 |
| rw_33_u | 3.196 | .030 | 106.066 | .000 | 3.506 |
| qw_25_u | 2.622 | .032 | 82.914 | .000 | 2.741 |
| qw_21_u | 3.061 | .027 | 114.221 | .000 | 3.777 |
| qw_13_u | 2.727 | .032 | 86.514 | .000 | 2.860 |
| gw_4_u | 3.537 | .028 | 127.605 | .000 | 4.222 |
| rw_18_u | 3.138 | .032 | 97.188 | .000 | 3.213 |
| gw_13_u | 3.132 | .038 | 83.334 | .000 | 2.756 |
| gw_3_u | 3.626 | .034 | 106.745 | .000 | 3.529 |
| sw_17_u | 3.741 | .032 | 116.320 | .000 | 3.865 |
| sw_2_u | 3.935 | .031 | 126.881 | .000 | 4.224 |
| sw_15_u | 4.002 | .032 | 125.139 | .000 | 4.144 |
| rw_36_u | 3.024 | .037 | 80.977 | .000 | 2.680 |
| qw_11 | 3.850 | .031 | 124.409 | .000 | 4.115 |
| rw_27 | 4.037 | .029 | 138.757 | .000 | 4.587 |
| gw_9 | 3.837 | .031 | 122.616 | .000 | 4.054 |
| gw_16 | 3.308 | .030 | 109.768 | .000 | 3.631 |
| rw_12 | 3.644 | .029 | 124.091 | .000 | 4.102 |
| rw_11 | 3.956 | .031 | 127.018 | .000 | 4.199 |
| sw_22 | 3.953 | .026 | 150.842 | .000 | 5.005 |
| ew_3 | 3.405 | .027 | 126.852 | .000 | 4.195 |
| ew_10 | 3.536 | .028 | 124.197 | .000 | 4.112 |
| ew_12 | 3.731 | .028 | 132.760 | .000 | 4.391 |
| ew_13 | 3.882 | .027 | 144.941 | .000 | 4.795 |
| ew_8 | 3.217 | .028 | 113.420 | .000 | 3.753 |

TABELLE E.5: Geschätzte Itemfehlervarianzen und Faktorvarianzen des finalen 26-Item vierfaktoriellen Modells mit korrelierten Itemfehlervarianzen ($N$ = 915)

| Items | geschätzte Varianzen | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| qw_26_u | .621 | .040 | 15.464 | .000 | .519 |
| rw_29_u | .640 | .036 | 17.767 | .000 | .641 |
| rw_33_u | .539 | .030 | 17.740 | .000 | .649 |
| qw_25_u | .464 | .030 | 15.736 | .000 | .507 |
| qw_21_u | .448 | .025 | 17.921 | .000 | .682 |
| qw_13_u | .572 | .032 | 18.141 | .000 | .629 |
| gw_4_u | .537 | .026 | 20.676 | .000 | .765 |
| rw_18_u | .716 | .036 | 20.129 | .000 | .751 |
| gw_13_u | .814 | .047 | 17.235 | .000 | .630 |
| gw_3_u | .884 | .042 | 21.121 | .000 | .837 |
| sw_17_u | .574 | .038 | 15.005 | .000 | .613 |
| sw_2_u | .523 | .035 | 14.829 | .000 | .603 |
| sw_15_u | .714 | .035 | 20.181 | .000 | .766 |
| rw_36_u | .925 | .047 | 19.516 | .000 | .726 |
| qw_11 | .451 | .036 | 12.596 | .000 | .515 |
| rw_27 | .538 | .036 | 15.140 | .000 | .694 |
| gw_9 | .705 | .036 | 19.546 | .000 | .786 |
| gw_16 | .738 | .035 | 20.798 | .000 | .889 |
| rw_12 | .576 | .031 | 18.385 | .000 | .730 |
| rw_11 | .683 | .041 | 16.469 | .000 | .770 |
| sw_22 | .501 | .029 | 17.515 | .000 | .803 |
| ew_3 | .346 | .032 | 10.928 | .000 | .526 |
| ew_10 | .557 | .034 | 16.461 | .000 | .753 |
| ew_12 | .553 | .029 | 18.933 | .000 | .766 |
| ew_13 | .480 | .028 | 17.091 | .000 | .733 |
| ew_8 | .518 | .032 | 15.963 | .000 | .705 |
| Faktoren | | | | | |
| Absolutheit | .576 | .049 | 11.815 | .000 | 1.000 |
| Einfachheit | .478 | .053 | 9.041 | .000 | 1.000 |
| Multimodalität | .425 | .041 | 10.302 | .000 | 1.000 |
| Entwicklung | .313 | .034 | 9.121 | .000 | 1.000 |

TABELLE E.6: Korrelierte Fehlervarianzen des vierfaktoriellen 26-Item Gesamtmodells sowie der vier einfaktoriellen Modelle der Gesamtstichprobe (*N* = 915)

| Modell | geschätzte Kovarianzen | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| vierfaktoriell Gesamtmodell (N = 915) | | | | | |
| rw_29_u mit rw_18_u | .129 | .030 | 4.353 | .000 | .190 |
| sw_17_u mit sw_22 | .097 | .023 | 4.223 | .000 | .181 |
| ew_12 mit ew_13 | .083 | .024 | 3.511 | .000 | .161 |
| ew_10 mit ew_8 | .086 | .025 | 3.422 | .001 | .161 |
| rw_29_u mit sw_2_u | -.097 | .022 | -4.321 | .000 | -.168 |
| einfaktorielles Absolutheit (N = 915) | | | | | |
| rw_29_u mit rw_18_u | .138 | .029 | 4.698 | .000 | .203 |
| qw_21_u mit rw_18_u | .072 | .023 | 3.151 | .002 | .127 |
| qw_26_u mit qw_21_u | -.069 | .021 | -3.260 | .001 | -.138 |
| einfaktorielles Einfachheit (N = 915) | | | | | |
| sw_2_u mit gw_13_u | -.108 | .039 | -2.776 | .005 | -.195 |
| gw_3_u | -.083 | .032 | -2.577 | .010 | -.144 |
| gw_13_u mit rw_36_u | .095 | .044 | 2.143 | .032 | .103 |
| einfaktorielles Multimodaltät (N = 915) | | | | | |
| qw_11 mit sw_22 | -.085 | .024 | -3.611 | .000 | -.202 |
| einfaktorielles Entwicklung (N = 915) | | | | | |
| ew_12 mit ew_13 | .089 | .025 | 3.566 | .000 | .171 |
| ew_10 mit ew_8 | .096 | .025 | 3.818 | .000 | .176 |

## Parameterschätzungen des finalen 26-Item vierfaktoriellen Gesamtmodells ($N$ = 915)

TABELLE E.7: Geschätzte semipartielle Regressionsgewichte (Ladungen $\lambda_{i1}$) der Items des finalen 26-Item vierfaktoriellen Modells ($N$ = 915)

| Factor regressed on item | geschätzte Ladungen | *SE* | *z* | *p* | Std.all |
|---|---|---|---|---|---|
| Absolutheit | | | | | |
| qw_26_u | 1.000 | | | | .691 |
| rw_29_u | .834 | .049 | 16.866 | .000 | .627 |
| rw_33_u | .703 | .046 | 15.125 | .000 | .583 |
| qw_25_u | .880 | .043 | 20.242 | .000 | .695 |
| qw_21_u | .609 | .046 | 13.364 | .000 | .568 |
| qw_13_u | .765 | .045 | 16.963 | .000 | .606 |
| gw_4_u | .542 | .043 | 12.462 | .000 | .489 |
| rw_18_u | .684 | .052 | 13.253 | .000 | .529 |
| Einfachheit | | | | | |
| gw_13_u | 1.000 | | | | .605 |
| gw_3_u | .600 | .064 | 9.401 | .000 | .401 |
| sw_17_u | .892 | .075 | 11.887 | .000 | .632 |
| sw_2_u | .879 | .073 | 12.048 | .000 | .645 |
| sw_15_u | .684 | .065 | 10.494 | .000 | .487 |
| rw_36_u | .852 | .072 | 11.850 | .000 | .520 |
| Multimodalität | | | | | |
| qw_11 | 1.000 | | | | .702 |
| rw_27 | .738 | .058 | 12.724 | .000 | .551 |
| gw_9 | .661 | .066 | 9.989 | .000 | .459 |
| gw_16 | .462 | .057 | 8.163 | .000 | .333 |
| rw_12 | .704 | .052 | 13.484 | .000 | .521 |
| rw_11 | .690 | .059 | 11.773 | .000 | .481 |
| sw_22 | .539 | .057 | 9.482 | .000 | .447 |
| Entwicklung | | | | | |
| ew_3 | 1.000 | | | | .651 |
| ew_10 | .877 | .083 | 10.571 | .000 | .539 |
| ew_12 | .833 | .077 | 10.860 | .000 | .518 |
| ew_13 | .833 | .075 | 11.174 | .000 | .544 |
| ew_8 | .931 | .069 | 13.530 | .000 | .574 |

TABELLE E.8: Geschätzte Kovarianzen zwischen den Faktoren des finalen 26-Item vierfaktoriellen Modells ($N$ = 915)

| Faktoren | geschätzte Kovarianzen | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| Absolutheit | | | | | |
| Einfachheit | .235 | .031 | 7.561 | .000 | .451 |
| Multimodalität | -.157 | .025 | -6.205 | .000 | -.317 |
| Entwicklung | -.094 | .022 | -4.373 | .000 | -.237 |
| Einfachheit | | | | | |
| Multimodalität | -.176 | .027 | -6.438 | .000 | -.389 |
| Entwicklung | -.004 | .018 | -0.221 | .825 | -.011 |
| Multimodalität | | | | | |
| Entwicklung | .136 | .018 | 7.661 | .000 | .392 |

TABELLE E.9: Geschätzte $R^2$ der Items des finalen 26-Item vierfaktoriellen Modells ($N$ = 915)

| Items | $R^2$ | Items | $R^2$ |
|---|---|---|---|
| qw_26_u | .477 | rw_36_u | .270 |
| rw_29_u | .393 | qw_11 | .493 |
| rw_33_u | .340 | rw_27 | .304 |
| qw_25_u | .483 | gw_9 | .211 |
| qw_21_u | .323 | gw_16 | .111 |
| qw_13_u | .367 | rw_12 | .271 |
| gw_4_u | .239 | rw_11 | .232 |
| rw_18_u | .280 | sw_22 | .200 |
| gw_13_u | .366 | ew_3 | .424 |
| gw_3_u | .161 | ew_10 | .290 |
| sw_17_u | .399 | ew_12 | .268 |
| sw_2_u | .416 | ew_13 | .296 |
| sw_15_u | .237 | ew_8 | .329 |

TABELLE E.10: Geschätzte Intercepts (Itemmittelwerte) der Items des finalen 26-Item vierfaktoriellen Modells ($N$ = 915)

| Items | geschätzte Intercepts | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| qw_26_u | 2.957 | .036 | 81.762 | .000 | 2.703 |
| rw_29_u | 3.077 | .033 | 92.509 | .000 | 3.059 |
| rw_33_u | 3.196 | .030 | 106.066 | .000 | 3.506 |
| qw_25_u | 2.622 | .032 | 82.914 | .000 | 2.741 |
| qw_21_u | 3.061 | .027 | 114.221 | .000 | 3.777 |
| qw_13_u | 2.727 | .032 | 86.514 | .000 | 2.860 |
| gw_4_u | 3.537 | .028 | 127.606 | .000 | 4.222 |
| rw_18_u | 3.138 | .032 | 97.188 | .000 | 3.213 |
| gw_13_u | 3.132 | .038 | 83.334 | .000 | 2.756 |
| gw_3_u | 3.626 | .034 | 106.745 | .000 | 3.529 |
| sw_17_u | 3.741 | .032 | 116.350 | .000 | 3.852 |
| sw_2_u | 3.935 | .031 | 126.872 | .000 | 4.197 |
| sw_15_u | 4.002 | .032 | 125.132 | .000 | 4.144 |
| rw_36_u | 3.024 | .037 | 80.974 | .000 | 2.680 |
| qw_11 | 3.850 | .031 | 124.408 | .000 | 4.115 |
| rw_27 | 4.037 | .029 | 138.757 | .000 | 4.587 |
| gw_9 | 3.837 | .031 | 122.615 | .000 | 4.054 |
| gw_16 | 3.308 | .030 | 109.767 | .000 | 3.631 |
| rw_12 | 3.644 | .029 | 124.091 | .000 | 4.102 |
| rw_11 | 3.956 | .031 | 127.017 | .000 | 4.199 |
| sw_22 | 3.953 | .026 | 150.842 | .000 | 4.987 |
| ew_3 | 3.405 | .027 | 126.850 | .000 | 4.195 |
| ew_10 | 3.536 | .028 | 124.159 | .000 | 4.111 |
| ew_12 | 3.731 | .028 | 132.762 | .000 | 4.391 |
| ew_13 | 3.882 | .027 | 144.960 | .000 | 4.796 |
| ew_8 | 3.217 | .028 | 113.418 | .000 | 3.753 |

TABELLE E.11: Geschätzte Itemfehlervarianzen und Faktorvarianzen des finalen 26-Item vierfaktoriellen Modells ($N$ = 915)

| Items | geschätzte Varianzen | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| qw_26_u | .626 | .040 | 15.587 | .000 | 0.523 |
| rw_29_u | .614 | .035 | 17.607 | .000 | 0.607 |
| rw_33_u | .548 | .031 | 17.890 | .000 | 0.660 |
| qw_25_u | .473 | .030 | 16.005 | .000 | 0.517 |
| qw_21_u | .445 | .025 | 17.928 | .000 | 0.677 |
| qw_13_u | .575 | .032 | 18.178 | .000 | 0.633 |
| gw_4_u | .534 | .026 | 20.745 | .000 | 0.761 |
| rw_18_u | .686 | .034 | 19.966 | .000 | 0.720 |
| gw_13_u | .818 | .049 | 16.863 | .000 | 0.634 |
| gw_3_u | .886 | .042 | 21.059 | .000 | 0.839 |
| sw_17_u | .567 | .039 | 14.555 | .000 | 0.601 |
| sw_2_u | .513 | .035 | 14.462 | .000 | 0.584 |
| sw_15_u | .712 | .036 | 20.042 | .000 | 0.763 |
| rw_36_u | .929 | .048 | 19.361 | .000 | 0.730 |
| qw_11 | .444 | .035 | 12.527 | .000 | 0.507 |
| rw_27 | .539 | .036 | 15.178 | .000 | 0.696 |
| gw_9 | .707 | .036 | 19.546 | .000 | 0.789 |
| gw_16 | .738 | .035 | 20.810 | .000 | 0.889 |
| rw_12 | .575 | .031 | 18.414 | .000 | 0.729 |
| rw_11 | .682 | .041 | 16.473 | .000 | 0.768 |
| sw_22 | .503 | .029 | 17.557 | .000 | 0.800 |
| ew_3 | .380 | .029 | 13.032 | .000 | 0.576 |
| ew_10 | .525 | .033 | 15.956 | .000 | 0.710 |
| ew_12 | .528 | .028 | 19.179 | .000 | 0.732 |
| ew_13 | .461 | .027 | 17.247 | .000 | 0.704 |
| ew_8 | .493 | .031 | 16.023 | .000 | 0.671 |
| Faktoren | | | | | |
| Absolutheit | .571 | .048 | 11.776 | .000 | 1.000 |
| Einfachheit | .473 | .054 | 8.766 | .000 | 1.000 |
| Multimodalität | .432 | .041 | 10.453 | .000 | 1.000 |
| Entwicklung | .279 | .030 | 9.190 | .000 | 1.000 |

## Parameterschätzungen des finalen 26-Item vierfaktoriellen Modells für Teilstichprobe 1

TABELLE E.12: Geschätzte semipartielle Regressionsgewichte (Ladungen $\lambda_{i1}$) der Items des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe S1 ($N$ = 457)

| Factor regressed on item | geschätzte Ladungen | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| Absolutheit S1 | | | | | |
| qw_26_u | 1.000 | | | | .669 |
| rw_29_u | .822 | .071 | 11.594 | .000 | .629 |
| rw_33_u | .703 | .062 | 11.406 | .000 | .588 |
| qw_25_u | .829 | .059 | 14.074 | .000 | .664 |
| qw_21_u | .543 | .060 | 8.981 | .000 | .527 |
| qw_13_u | .773 | .069 | 11.243 | .000 | .615 |
| gw_4_u | .570 | .064 | 8.950 | .000 | .528 |
| rw_18_u | .681 | .075 | 9.037 | .000 | .545 |
| Einfachheit S1 | | | | | |
| gw_13_u | 1.000 | | | | .597 |
| gw_3_u | .569 | .088 | 6.466 | .000 | .379 |
| sw_17_u | .928 | .114 | 8.125 | .000 | .668 |
| sw_2_u | .980 | .116 | 8.470 | .000 | .685 |
| sw_15_u | .660 | .091 | 7.245 | .000 | .474 |
| rw_36_u | .816 | .093 | 8.737 | .000 | .511 |
| Multimodalität S1 | | | | | |
| qw_11 | 1.000 | | | | .657 |
| rw_27 | .831 | .086 | 9.659 | .000 | .606 |
| gw_9 | .816 | .100 | 8.184 | .000 | .533 |
| gw_16 | .550 | .083 | 6.646 | .000 | .369 |
| rw_12 | .744 | .080 | 9.351 | .000 | .541 |
| rw_11 | .700 | .091 | 7.729 | .000 | .467 |
| sw_22 | .539 | .087 | 6.179 | .000 | .429 |
| Entwicklung S1 | | | | | |
| ew_10 | 1.000 | | | | .589 |
| ew_3 | .980 | .123 | 7.998 | .000 | .625 |
| ew_12 | .798 | .119 | 6.701 | .000 | .493 |
| ew_13 | .750 | .130 | 5.790 | .000 | .473 |
| ew_8 | .905 | .105 | 8.628 | .000 | .550 |

TABELLE E.13: Geschätzte Kovarianzen zwischen den Faktoren des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe S1 ($N$ = 457)

| Faktoren | geschätzte Kovarianzen | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| Absolutheit S1 | | | | | |
| Einfachheit S1 | .264 | .048 | 5.526 | .000 | .511 |
| Multimodalität S1 | -.174 | .039 | -4.479 | .000 | -.368 |
| Entwicklung S1 | -.055 | .028 | -1.936 | .053 | -.144 |
| Einfachheit S1 | | | | | |
| Multimodalität S1 | -.195 | .040 | -4.902 | .000 | -.452 |
| Entwicklung S1 | -.013 | .024 | -.564 | .573 | -.038 |
| Multimodalität S1 | | | | | |
| Entwicklung S1 | .132 | .031 | 4.218 | .000 | .415 |

TABELLE E.14: Geschätzte $R^2$ der Items des finalen 26-Item vierfaktoriellen Modells von Teilstichprobe S1 ($N$ = 457)

| Items | $R^2$ | Items | $R^2$ |
|---|---|---|---|
| qw_26_u | .448 | rw_36_u | .261 |
| rw_29_u | .395 | qw_11 | .432 |
| rw_33_u | .345 | rw_27 | .367 |
| qw_25_u | .440 | gw_9 | .284 |
| qw_21_u | .278 | gw_16 | .136 |
| qw_13_u | .378 | rw_12 | .293 |
| gw_4_u | .279 | rw_11 | .218 |
| rw_18_u | .297 | sw_22 | .184 |
| gw_13_u | .356 | ew_10 | .347 |
| gw_3_u | .143 | ew_3 | .390 |
| sw_17_u | .446 | ew_12 | .243 |
| sw_2_u | .470 | ew_13 | .224 |
| sw_15_u | .225 | ew_8 | .302 |

TABELLE E.15: Geschätzte Intercepts (Itemmittelwerte) der Items des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe S1 ($N$ = 457)

| Items | geschätzte Intercepts | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| qw_26_u | 2.932 | .053 | 55.693 | .000 | 2.605 |
| rw_29_u | 3.081 | .046 | 66.886 | .000 | 3.131 |
| rw_33_u | 3.127 | .042 | 74.226 | .000 | 3.472 |
| qw_25_u | 2.608 | .044 | 59.239 | .000 | 2.771 |
| qw_21_u | 3.033 | .036 | 83.545 | .000 | 3.910 |
| qw_13_u | 2.685 | .044 | 60.551 | .000 | 2.832 |
| gw_4_u | 3.519 | .038 | 92.316 | .000 | 4.324 |
| rw_18_u | 3.109 | .044 | 70.651 | .000 | 3.305 |
| gw_13_u | 3.105 | .054 | 57.693 | .000 | 2.699 |
| gw_3_u | 3.630 | .048 | 75.215 | .000 | 3.518 |
| sw_17_u | 3.704 | .045 | 82.832 | .000 | 3.882 |
| sw_2_u | 3.912 | .046 | 85.155 | .000 | 3.988 |
| sw_15_u | 3.988 | .045 | 89.110 | .000 | 4.177 |
| rw_36_u | 2.968 | .051 | 57.806 | .000 | 2.709 |
| qw_11 | 3.821 | .045 | 85.473 | .000 | 3.998 |
| rw_27 | 4.068 | .040 | 100.982 | .000 | 4.724 |
| gw_9 | 3.808 | .045 | 84.694 | .000 | 3.963 |
| gw_16 | 3.290 | .044 | 75.036 | .000 | 3.514 |
| rw_12 | 3.650 | .040 | 90.462 | .000 | 4.232 |
| rw_11 | 3.958 | .044 | 89.978 | .000 | 4.209 |
| sw_22 | 3.954 | .037 | 107.219 | .000 | 5.016 |
| ew_10 | 3.564 | .040 | 88.346 | .000 | 4.145 |
| ew_3 | 3.389 | .037 | 91.170 | .000 | 4.265 |
| ew_12 | 3.723 | .038 | 96.884 | .000 | 4.537 |
| ew_13 | 3.849 | .038 | 102.264 | .000 | 4.792 |
| ew_8 | 3.250 | .039 | 83.184 | .000 | 3.895 |

TABELLE E.16: Geschätzte Itemfehlervarianzen und Faktorvarianzen des finalen 26-Item vierfaktoriellen Modells von Teilstichprobe S1 ($N$ = 457)

| Items | geschätzte Varianzen | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| qw_26_u | .699 | .058 | 11.980 | .000 | .552 |
| rw_29_u | .586 | .046 | 12.606 | .000 | .605 |
| rw_33_u | .531 | .043 | 12.388 | .000 | .655 |
| qw_25_u | .496 | .039 | 12.747 | .000 | .560 |
| qw_21_u | .435 | .035 | 12.268 | .000 | .722 |
| qw_13_u | .559 | .046 | 12.283 | .000 | .622 |
| gw_4_u | .478 | .033 | 14.597 | .000 | .721 |
| rw_18_u | .622 | .046 | 13.461 | .000 | .703 |
| gw_13_u | .853 | .070 | 12.249 | .000 | .644 |
| gw_3_u | .912 | .063 | 14.521 | .000 | .857 |
| sw_17_u | .505 | .051 | 9.823 | .000 | .554 |
| sw_2_u | .510 | .055 | 9.352 | .000 | .530 |
| sw_15_u | .707 | .046 | 15.342 | .000 | .775 |
| rw_36_u | .887 | .066 | 13.489 | .000 | .739 |
| qw_11 | .519 | .054 | 9.618 | .000 | .568 |
| rw_27 | .469 | .042 | 11.270 | .000 | .633 |
| gw_9 | .661 | .050 | 13.300 | .000 | .716 |
| gw_16 | .757 | .055 | 13.887 | .000 | .864 |
| rw_12 | .526 | .043 | 12.120 | .000 | .707 |
| rw_11 | .691 | .054 | 12.763 | .000 | .782 |
| sw_22 | .507 | .044 | 11.408 | .000 | .816 |
| ew_10 | .482 | .049 | 9.872 | .000 | .653 |
| ew_3 | .385 | .042 | 9.175 | .000 | .610 |
| ew_12 | .510 | .037 | 13.605 | .000 | .757 |
| ew_13 | .501 | .038 | 13.206 | .000 | .776 |
| ew_8 | .486 | .040 | 12.205 | .000 | .698 |
| Faktoren | | | | | |
| Absolutheit S1 | .567 | .071 | 7.993 | .000 | 1.000 |
| Einfachheit S1 | .471 | .079 | 5.959 | .000 | 1.000 |
| Multimodalität S1 | .394 | .056 | 7.061 | .000 | 1.000 |
| Entwicklung S1 | .257 | .050 | 5.145 | .000 | 1.000 |

## Parameterschätzungen des finalen 26-Item vierfaktoriellen Modells für Teilstichprobe 2

TABELLE E.17: Geschätzte semipartielle Regressionsgewichte (Ladungen $\lambda_{i1}$) der Items des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe S2 ($N$ = 458)

| Factor regressed on item | geschätzte Ladungen | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| Absolutheit S2 | | | | | |
| qw_26_u | 1.000 | | | | .715 |
| rw_29_u | .850 | .069 | 12.397 | .000 | .629 |
| rw_33_u | .691 | .070 | 9.848 | .000 | .572 |
| qw_25_u | .926 | .065 | 14.306 | .000 | .724 |
| qw_21_u | .669 | .069 | 9.746 | .000 | .602 |
| qw_13_u | .759 | .060 | 12.683 | .000 | .602 |
| gw_4_u | .515 | .061 | 8.448 | .000 | .454 |
| rw_18_u | .687 | .072 | 9.593 | .000 | .516 |
| Einfachheit S2 | | | | | |
| gw_13_u | 1.000 | | | | .613 |
| gw_3_u | .634 | .095 | 6.678 | .000 | .425 |
| sw_17_u | .860 | .101 | 8.540 | .000 | .599 |
| sw_2_u | .787 | .096 | 8.159 | .000 | .607 |
| sw_15_u | .717 | .095 | 7.509 | .000 | .504 |
| rw_36_u | .884 | .113 | 7.814 | .000 | .525 |
| Multimodalität S2 | | | | | |
| qw_11 | 1.000 | | | | .763 |
| rw_27 | .645 | .078 | 8.225 | .000 | .502 |
| gw_9 | .500 | .087 | 5.762 | .000 | .375 |
| gw_16 | .358 | .076 | 4.692 | .000 | .283 |
| rw_12 | .675 | .068 | 9.909 | .000 | .516 |
| rw_11 | .670 | .076 | 8.810 | .000 | .495 |
| sw_22 | .530 | .075 | 7.034 | .000 | .464 |
| Entwicklung S2 | | | | | |
| ew_3 | 1.000 | | | | .676 |
| ew_12 | .855 | .113 | 7.567 | .000 | .545 |
| ew_13 | .883 | .107 | 8.260 | .000 | .607 |
| ew_8 | .917 | .082 | 11.156 | .000 | .584 |
| ew_10 | .764 | .105 | 7.275 | .000 | .497 |

TABELLE E.18: Geschätzte Kovarianzen zwischen den Faktoren des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe S2 ($N$ = 458)

| Faktoren | Geschätzte Kovarianzen | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| Absolutheit S2 | | | | | |
| Einfachheit S2 | .197 | .042 | 4.663 | .000 | .378 |
| Multimodalität S2 | -.136 | .034 | -3.998 | .000 | -.256 |
| Entwicklung S2 | -.133 | .032 | -4.206 | .000 | -.313 |
| Einfachheit S2 | | | | | |
| Multimodalität S2 | -.143 | .039 | -3.653 | .000 | -.299 |
| Entwicklung S2 | .007 | .027 | .266 | .790 | .019 |
| Multimodalität S2 | | | | | |
| Entwicklung S2 | .140 | .026 | 5.388 | .000 | .357 |

TABELLE E.19: Geschätzte $R^2$ der Items des finalen 26-Item vierfaktoriellen Modells von Teilstichprobe S2 ($N$ = 458)

| Items | $R^2$ | Items | $R^2$ |
|---|---|---|---|
| qw_26_u | .512 | rw_36_u | .276 |
| rw_29_u | .395 | qw_11 | .583 |
| rw_33_u | .328 | rw_27 | .252 |
| qw_25_u | .524 | gw_9 | .140 |
| qw_21_u | .363 | gw_16 | .080 |
| qw_13_u | .362 | rw_12 | .266 |
| gw_4_u | .206 | rw_11 | .246 |
| rw_18_u | .266 | sw_22 | .215 |
| gw_13_u | .375 | ew_10 | .247 |
| gw_3_u | .181 | ew_3 | .457 |
| sw_17_u | .359 | ew_12 | .297 |
| sw_2_u | .368 | ew_13 | .368 |
| sw_15_u | .254 | ew_8 | .342 |

TABELLE E.20: Geschätzte Intercepts (Itemmittelwerte) der Items des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe S2 ($N$ = 458)

| Items | geschätzte Intercepts | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| .qw_26_u | 2.983 | .050 | 60.143 | .000 | 2.810 |
| .rw_29_u | 3.072 | .048 | 64.046 | .000 | 2.993 |
| .rw_33_u | 3.264 | .043 | 76.189 | .000 | 3.560 |
| .qw_25_u | 2.635 | .045 | 58.065 | .000 | 2.713 |
| .qw_21_u | 3.090 | .039 | 78.450 | .000 | 3.666 |
| .qw_13_u | 2.769 | .045 | 61.911 | .000 | 2.893 |
| .gw_4_u | 3.555 | .040 | 88.403 | .000 | 4.131 |
| .rw_18_u | 3.166 | .047 | 67.070 | .000 | 3.134 |
| .gw_13_u | 3.160 | .052 | 60.229 | .000 | 2.817 |
| .gw_3_u | 3.622 | .048 | 75.750 | .000 | 3.540 |
| .sw_17_u | 3.778 | .046 | 81.841 | .000 | 3.828 |
| .sw_2_u | 3.959 | .042 | 94.999 | .000 | 4.439 |
| .sw_15_u | 4.017 | .046 | 87.904 | .000 | 4.113 |
| .rw_36_u | 3.080 | .054 | 56.909 | .000 | 2.661 |
| .qw_11 | 3.879 | .043 | 90.727 | .000 | 4.243 |
| .rw_27 | 4.007 | .042 | 95.529 | .000 | 4.464 |
| .gw_9 | 3.867 | .043 | 88.898 | .000 | 4.154 |
| .gw_16 | 3.325 | .041 | 80.434 | .000 | 3.758 |
| .rw_12 | 3.638 | .043 | 85.259 | .000 | 3.984 |
| .rw_11 | 3.954 | .044 | 89.654 | .000 | 4.189 |
| .sw_22 | 3.952 | .037 | 106.114 | .000 | 4.958 |
| .ew_3 | 3.421 | .039 | 88.358 | .000 | 4.132 |
| .ew_12 | 3.738 | .041 | 91.178 | .000 | 4.260 |
| .ew_13 | 3.915 | .038 | 102.890 | .000 | 4.808 |
| .ew_8 | 3.184 | .041 | 77.524 | .000 | 3.626 |
| .ew_10 | 3.509 | .040 | 87.349 | .000 | 4.082 |

TABELLE E.21: Geschätzte Itemfehlervarianzen und Faktorvarianzen des finalen 26-Item vierfaktoriellen Modells von Teilstichprobe S2 ($N$ = 458)

| Items | geschätzte Varianzen | *SE* | *z* | *p* | vollst. std |
|---|---|---|---|---|---|
| qw_26_u | .550 | .055 | 10.032 | .000 | .488 |
| rw_29_u | .637 | .051 | 12.427 | .000 | .605 |
| rw_33_u | .565 | .044 | 12.785 | .000 | .672 |
| qw_25_u | .450 | .044 | 10.228 | .000 | .476 |
| qw_21_u | .453 | .035 | 13.035 | .000 | .637 |
| qw_13_u | .584 | .044 | 13.348 | .000 | .638 |
| gw_4_u | .588 | .040 | 14.788 | .000 | .794 |
| rw_18_u | .749 | .051 | 14.675 | .000 | .734 |
| gw_13_u | .786 | .069 | 11.452 | .000 | .625 |
| gw_3_u | .858 | .056 | 15.281 | .000 | .819 |
| sw_17_u | .624 | .059 | 10.532 | .000 | .641 |
| sw_2_u | .503 | .048 | 10.472 | .000 | .632 |
| sw_15_u | .711 | .054 | 13.071 | .000 | .746 |
| rw_36_u | .970 | .073 | 13.295 | .000 | .724 |
| qw_11 | .349 | .044 | 7.950 | .000 | .417 |
| rw_27 | .603 | .054 | 11.081 | .000 | .748 |
| gw_9 | .745 | .052 | 14.255 | .000 | .860 |
| gw_16 | .720 | .046 | 15.724 | .000 | .920 |
| rw_12 | .612 | .044 | 13.945 | .000 | .734 |
| rw_11 | .672 | .061 | 10.963 | .000 | .754 |
| sw_22 | .499 | .036 | 14.020 | .000 | .785 |
| ew_3 | .372 | .041 | 9.083 | .000 | .543 |
| ew_12 | .541 | .040 | 13.416 | .000 | .703 |
| ew_13 | .419 | .038 | 11.147 | .000 | .632 |
| ew_8 | .508 | .049 | 10.299 | .000 | .658 |
| ew_10 | .556 | .043 | 12.860 | .000 | .753 |
| Faktoren | | | | | |
| Absolutheit S2 | .576 | .067 | 8.609 | .000 | 1.000 |
| Einfachheit S2 | .472 | .075 | 6.334 | .000 | 1.000 |
| Multimodalität S2 | .487 | .065 | 7.499 | .000 | 1.000 |
| Entwicklung S2 | .313 | .047 | 6.696 | .000 | 1.000 |

# F Itemkorrelations- und Residualkovarianzmatrizen Studie 2

TABELLE F.1: Itemkorrelationsmatrix der Gesamtstichprobe ($N$ = 915) – Teil 1

| Items | gw_17 | rw_33_u | gw_3_u | ew_5_u | rw_11 | sw_2_u | qw_21_u | ew_13 | sw_11_u | rw_18_u | rw_27 | ew_8 | rw_8_u | sw_17_u | sw_22 | sw_15_u | rw_12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| gw_17 | – | | | | | | | | | | | | | | | | |
| rw_33_u | .05 | – | | | | | | | | | | | | | | | |
| gw_3_u | -.05 | .28 | – | | | | | | | | | | | | | | |
| ew_5_u | .00 | .19 | .15 | – | | | | | | | | | | | | | |
| rw_11 | .10 | -.18 | -.03 | -.01 | – | | | | | | | | | | | | |
| sw_2_u | -.13 | .26 | .21 | .14 | -.11 | – | | | | | | | | | | | |
| qw_21_u | .06 | .36 | .21 | .19 | -.12 | .10 | – | | | | | | | | | | |
| ew_13 | .09 | -.01 | -.02 | -.19 | .13 | -.01 | -.15 | – | | | | | | | | | |
| sw_11_u | -.02 | .04 | .08 | .08 | .00 | .21 | .04 | .01 | – | | | | | | | | |
| rw_18_u | .05 | .24 | .17 | .20 | -.08 | .11 | .38 | -.07 | .02 | – | | | | | | | |
| rw_27 | .24 | -.24 | -.09 | -.11 | .22 | -.17 | -.14 | .12 | -.03 | -.02 | – | | | | | | |
| ew_8 | .02 | -.02 | .00 | -.10 | .11 | .02 | -.10 | .27 | .02 | -.10 | .11 | – | | | | | |
| rw_8_u | .12 | .21 | .13 | .12 | .01 | -.02 | .26 | .05 | .02 | .26 | .05 | -.01 | – | | | | |
| sw_17_u | -.13 | .25 | .24 | .13 | -.08 | .47 | .16 | -.05 | .20 | .16 | -.15 | .00 | .01 | – | | | |
| sw_22 | .18 | -.07 | -.04 | -.04 | .20 | -.02 | -.03 | .16 | .02 | -.04 | .31 | .03 | .05 | .05 | – | | |
| sw_15_u | -.12 | .14 | .17 | .12 | -.10 | .38 | .00 | .05 | .24 | .07 | -.14 | .10 | -.10 | .30 | -.09 | – | |
| rw_12 | .13 | -.14 | -.07 | -.04 | .25 | -.06 | -.09 | .09 | .02 | -.03 | .29 | .05 | .00 | -.04 | .29 | -.10 | – |
| ew_3 | .04 | -.07 | .06 | -.11 | .14 | .00 | -.09 | .38 | .03 | -.12 | .12 | .40 | -.02 | .01 | .10 | .12 | .10 |
| qw_25_u | -.04 | .41 | .14 | .19 | -.18 | .13 | .39 | -.13 | .04 | .32 | -.20 | -.10 | .25 | .22 | -.06 | .04 | -.08 |
| rw_36_u | -.14 | .23 | .22 | .21 | -.09 | .32 | .21 | -.12 | .11 | .21 | -.17 | -.03 | .04 | .29 | -.07 | .20 | -.06 |
| ew_10 | .14 | -.01 | -.01 | -.15 | .18 | -.04 | -.05 | .25 | -.04 | -.02 | .21 | .39 | .04 | .00 | .14 | -.02 | .20 |
| gw_9 | .16 | -.17 | -.09 | -.09 | .21 | -.15 | -.07 | .13 | .01 | -.06 | .26 | .06 | .03 | -.15 | .23 | -.13 | .18 |
| qw_26_u | -.01 | .39 | .15 | .21 | -.14 | .12 | .34 | -.10 | .00 | .33 | -.15 | -.08 | .27 | .19 | -.07 | .06 | -.04 |
| ew_12 | .07 | -.07 | -.01 | -.16 | .07 | .03 | -.14 | .37 | .05 | -.06 | .10 | .26 | -.02 | .06 | .12 | .13 | .02 |
| qw_11 | .22 | -.21 | -.07 | -.08 | .35 | -.19 | -.11 | .14 | .04 | -.05 | .37 | .12 | .01 | -.11 | .28 | -.15 | .41 |
| sw_3_u | -.17 | -.04 | .08 | .09 | .00 | .13 | -.05 | -.05 | .17 | -.01 | -.08 | .08 | -.17 | .14 | .00 | .24 | -.06 |
| qw_13_u | -.09 | .32 | .23 | .25 | -.11 | .17 | .31 | -.11 | .07 | .30 | -.14 | -.07 | .18 | .24 | -.05 | .09 | .02 |
| sw_9 | .06 | .06 | .06 | .03 | .07 | .11 | .04 | .01 | .11 | .04 | .07 | .05 | .01 | .15 | .20 | .07 | .12 |
| gw_13_u | -.21 | .25 | .25 | .19 | -.16 | .35 | .12 | -.11 | .10 | .15 | -.21 | -.01 | -.07 | .36 | -.17 | .30 | -.12 |
| ew_2 | .15 | -.13 | -.15 | -.12 | .08 | -.09 | -.17 | .20 | .02 | -.09 | .16 | .14 | .04 | -.13 | .09 | -.07 | .06 |
| gw_4_u | .10 | .25 | .06 | .18 | -.07 | .05 | .28 | -.01 | .08 | .27 | .02 | -.07 | .22 | .13 | .07 | .01 | .02 |
| gw_16 | .11 | -.14 | .00 | .08 | .17 | -.10 | .00 | .03 | -.02 | .00 | .13 | .04 | -.03 | -.06 | .17 | -.06 | .15 |
| rw_29_u | .00 | .32 | .18 | .14 | -.11 | .00 | .36 | -.06 | -.04 | .45 | -.10 | -.07 | .27 | .08 | -.07 | -.01 | -.06 |
| ew_16_u | -.03 | .10 | .09 | .22 | -.03 | .08 | .19 | -.25 | .00 | .18 | -.06 | -.20 | .13 | .11 | -.02 | -.01 | .01 |

TABELLE F.2: Itemkorrelationsmatrix der Gesamtstichprobe ($N$ = 915) – Teil 2

| Items | ew_3 | qw_25_u | rw_36_u | ew_10 | gw_9 | qw_26_u | ew_12 | qw_11 | sw_3_u | qw_13_u | sw_9 | gw_13_u | ew_2 | gw_4_u | gw_16 | rw_29_u |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| gw_17 | | | | | | | | | | | | | | | | |
| rw_33_u | | | | | | | | | | | | | | | | |
| gw_3_u | | | | | | | | | | | | | | | | |
| ew_5_u | | | | | | | | | | | | | | | | |
| rw_11 | | | | | | | | | | | | | | | | |
| sw_2_u | | | | | | | | | | | | | | | | |
| qw_21_u | | | | | | | | | | | | | | | | |
| ew_13 | | | | | | | | | | | | | | | | |
| sw_11_u | | | | | | | | | | | | | | | | |
| rw_18_u | | | | | | | | | | | | | | | | |
| rw_27 | | | | | | | | | | | | | | | | |
| ew_8 | | | | | | | | | | | | | | | | |
| rw_8_u | | | | | | | | | | | | | | | | |
| sw_17_u | | | | | | | | | | | | | | | | |
| sw_22 | | | | | | | | | | | | | | | | |
| sw_15_u | | | | | | | | | | | | | | | | |
| rw_12 | | | | | | | | | | | | | | | | |
| ew_3 | – | | | | | | | | | | | | | | | |
| qw_25_u | -.17 | – | | | | | | | | | | | | | | |
| rw_36_u | -.04 | .24 | – | | | | | | | | | | | | | |
| ew_10 | .33 | -.09 | -.06 | – | | | | | | | | | | | | |
| gw_9 | .14 | -.15 | -.17 | .15 | – | | | | | | | | | | | |
| qw_26_u | -.12 | **.52** | .26 | -.03 | -.16 | – | | | | | | | | | | |
| ew_12 | .31 | -.18 | -.08 | .23 | .13 | -.15 | – | | | | | | | | | |
| qw_11 | .12 | -.18 | -.14 | .28 | .31 | -.15 | .15 | – | | | | | | | | |
| sw_3_u | .05 | -.09 | .13 | -.05 | -.06 | -.10 | .07 | -.06 | – | | | | | | | |
| qw_13_u | -.09 | .40 | .33 | -.08 | -.12 | .43 | -.14 | -.15 | .04 | – | | | | | | |
| sw_9 | .07 | .05 | .04 | .11 | .07 | .01 | .06 | .14 | .04 | .10 | – | | | | | |
| gw_13_u | .01 | .19 | .34 | -.08 | -.19 | .18 | -.03 | -.19 | .21 | .30 | .07 | – | | | | |
| ew_2 | .18 | -.18 | -.08 | .20 | .15 | -.14 | .26 | .16 | -.04 | -.20 | -.01 | -.22 | – | | | |
| gw_4_u | -.09 | .33 | .13 | .05 | .01 | .30 | -.02 | -.01 | -.07 | .29 | .07 | .08 | .08 | – | | |
| gw_16 | .05 | -.09 | -.04 | .14 | .20 | -.07 | .08 | .24 | .00 | -.03 | .22 | -.12 | .08 | .05 | – | |
| rw_29_u | -.13 | .39 | .21 | -.02 | -.10 | .45 | -.08 | -.09 | -.08 | .37 | .01 | .11 | -.15 | .35 | .00 | – |
| ew_16_u | -.21 | .24 | .15 | -.14 | -.07 | .23 | **-.34** | -.03 | .03 | .22 | .05 | .13 | -.16 | .09 | .05 | .21 |

Die stärkste negative und positive Korrelation ist hervorgehoben.

TABELLE F.3: Itemkorrelationsmatrix der Teilstichprobe S1 ($N$ = 457) – Teil 1

| Items | gw_17 | rw_33_u | gw_3_u | ew_5_u | rw_11 | sw_2_u | qw_21_u | ew_13 | sw_11_u | rw_18_u | rw_27 | ew_8 | rw_8_u | sw_17_u | sw_22 | sw_15_u | rw_12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| gw_17 | – | | | | | | | | | | | | | | | | |
| rw_33_u | .06 | – | | | | | | | | | | | | | | | |
| gw_3_u | -.11 | .31 | – | | | | | | | | | | | | | | |
| ew_5_u | .00 | .21 | .19 | – | | | | | | | | | | | | | |
| rw_11 | .09 | -.16 | -.04 | -.04 | – | | | | | | | | | | | | |
| sw_2_u | -.15 | .23 | .20 | .14 | -.14 | – | | | | | | | | | | | |
| qw_21_u | .11 | .35 | .21 | .15 | -.05 | .10 | – | | | | | | | | | | |
| ew_13 | .10 | -.05 | -.05 | -.13 | .10 | -.08 | -.09 | – | | | | | | | | | |
| sw_11_u | -.04 | .10 | .11 | .15 | .04 | .21 | .06 | -.02 | – | | | | | | | | |
| rw_18_u | .07 | .28 | .19 | .20 | -.06 | .13 | .39 | -.04 | .06 | – | | | | | | | |
| rw_27 | .22 | -.27 | -.10 | -.14 | .26 | -.19 | -.16 | .14 | -.03 | -.02 | – | | | | | | |
| ew_8 | .00 | .01 | .03 | -.04 | .10 | .03 | -.14 | .26 | -.01 | -.10 | .12 | – | | | | | |
| rw_8_u | .20 | .21 | .11 | .15 | .04 | .01 | .26 | .09 | .04 | .28 | .09 | -.05 | – | | | | |
| sw_17_u | -.16 | .22 | .22 | .17 | -.09 | **.54** | .13 | -.09 | .25 | .19 | -.17 | .02 | -.02 | – | | | |
| sw_22 | .25 | -.12 | -.04 | -.04 | .20 | .00 | -.03 | .14 | .00 | -.01 | .35 | .03 | .12 | .04 | – | | |
| sw_15_u | -.16 | .13 | .17 | .16 | -.16 | .37 | .03 | .01 | .22 | .12 | -.17 | .14 | -.07 | .32 | -.11 | – | |
| rw_12 | .22 | -.17 | -.11 | -.07 | .24 | -.07 | -.07 | .07 | .04 | -.03 | .34 | .06 | .05 | -.08 | .28 | -.12 | – |
| ew_3 | .05 | -.08 | .12 | -.08 | .11 | .00 | -.09 | .33 | .01 | -.06 | .16 | .33 | -.05 | -.03 | .08 | .12 | .15 |
| qw_25_u | -.02 | .38 | .11 | .20 | -.15 | .17 | .35 | -.07 | .05 | .32 | -.20 | -.11 | .25 | .21 | -.04 | .05 | -.10 |
| rw_36_u | -.11 | .23 | .17 | .24 | -.13 | .32 | .18 | -.12 | .10 | .21 | -.19 | -.05 | .05 | .28 | -.07 | .18 | -.14 |
| ew_10 | .17 | .04 | -.02 | -.13 | .15 | -.03 | -.06 | .22 | .00 | .00 | .24 | .38 | .09 | .00 | .15 | -.04 | .20 |
| gw_9 | .19 | -.17 | -.10 | -.11 | .23 | -.17 | -.02 | .11 | .04 | -.07 | .32 | .08 | .12 | -.16 | .24 | -.19 | .27 |
| qw_26_u | -.01 | .39 | .14 | .24 | -.15 | .18 | .29 | -.06 | .01 | .29 | -.23 | -.07 | .21 | .22 | -.06 | .14 | -.09 |
| ew_12 | .03 | -.06 | -.02 | -.18 | .02 | -.05 | -.08 | .31 | -.01 | -.03 | .10 | .24 | -.02 | .05 | .10 | .13 | .00 |
| qw_11 | .22 | -.22 | -.12 | -.15 | .31 | -.22 | -.08 | .17 | .07 | -.08 | .35 | .15 | .07 | -.12 | .23 | -.18 | .38 |
| sw_3_u | -.25 | -.05 | .07 | .17 | .02 | .11 | -.10 | -.06 | .15 | -.04 | -.11 | .09 | -.19 | .11 | -.01 | .25 | -.06 |
| qw_13_u | -.08 | .31 | .22 | .25 | -.09 | .23 | .27 | -.11 | .17 | .30 | -.13 | -.02 | .10 | .29 | -.07 | .17 | .02 |
| sw_9 | .11 | .08 | .08 | .07 | .06 | .13 | .05 | -.07 | .07 | .05 | .08 | .01 | .03 | .20 | .24 | .07 | .11 |
| gw_13_u | -.25 | .24 | .24 | .24 | -.16 | .38 | .08 | -.16 | .09 | .16 | -.25 | -.06 | -.07 | .35 | -.15 | .26 | -.15 |
| ew_2 | .18 | -.13 | -.19 | -.17 | .11 | -.13 | -.15 | .20 | -.02 | -.14 | .16 | .11 | .07 | -.14 | .07 | -.11 | .10 |
| gw_4_u | .16 | .23 | .05 | .14 | -.06 | .06 | .29 | .05 | .05 | .29 | .00 | -.07 | .25 | .16 | .08 | .01 | .06 |
| gw_16 | .12 | -.18 | -.02 | .09 | .19 | -.10 | -.02 | -.04 | .03 | -.03 | .12 | .04 | .04 | -.05 | .19 | -.04 | .20 |
| rw_29_u | .05 | .34 | .16 | .14 | -.11 | .05 | .32 | -.03 | .00 | .44 | -.14 | -.07 | .23 | .07 | -.05 | .03 | -.15 |
| ew_16_u | .05 | .09 | .05 | .22 | -.02 | .09 | .14 | -.24 | .00 | .13 | -.08 | -.18 | .12 | .07 | -.01 | .01 | .01 |

Die stärkste negative und positive Korrelation ist hervorgehoben.

TABELLE F.4: Itemkorrelationsmatrix der Teilstichprobe S1 ($N$ = 457) – Teil 2

| Items | ew_3 | qw_25_u | rw_36_u | ew_10 | gw_9 | qw_26_u | ew_12 | qw_11 | sw_3_u | qw_13_u | sw_9 | gw_13_u | ew_2 | gw_4_u | gw_16 | rw_29_u | ew_16_u |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| gw_17 | | | | | | | | | | | | | | | | | |
| rw_33_u | | | | | | | | | | | | | | | | | |
| gw_3_u | | | | | | | | | | | | | | | | | |
| ew_5_u | | | | | | | | | | | | | | | | | |
| rw_11 | | | | | | | | | | | | | | | | | |
| sw_2_u | | | | | | | | | | | | | | | | | |
| qw_21_u | | | | | | | | | | | | | | | | | |
| ew_13 | | | | | | | | | | | | | | | | | |
| sw_11_u | | | | | | | | | | | | | | | | | |
| rw_18_u | | | | | | | | | | | | | | | | | |
| rw_27 | | | | | | | | | | | | | | | | | |
| ew_8 | | | | | | | | | | | | | | | | | |
| rw_8_u | | | | | | | | | | | | | | | | | |
| sw_17_u | | | | | | | | | | | | | | | | | |
| sw_22 | | | | | | | | | | | | | | | | | |
| sw_15_u | | | | | | | | | | | | | | | | | |
| rw_12 | | | | | | | | | | | | | | | | | |
| ew_3 | – | | | | | | | | | | | | | | | | |
| qw_25_u | -.12 | – | | | | | | | | | | | | | | | |
| rw_36_u | -.03 | .28 | – | | | | | | | | | | | | | | |
| ew_10 | .35 | -.11 | -.08 | – | | | | | | | | | | | | | |
| gw_9 | .17 | -.14 | -.17 | .11 | – | | | | | | | | | | | | |
| qw_26_u | -.07 | .49 | .30 | -.02 | -.21 | – | | | | | | | | | | | |
| ew_12 | .30 | -.08 | -.10 | .24 | .12 | -.09 | – | | | | | | | | | | |
| qw_11 | .13 | -.18 | -.22 | .30 | .34 | -.20 | .19 | – | | | | | | | | | |
| sw_3_u | .04 | -.09 | .10 | -.06 | -.06 | -.02 | .04 | -.06 | – | | | | | | | | |
| qw_13_u | -.02 | .39 | .38 | -.03 | -.11 | .37 | -.05 | -.21 | .07 | – | | | | | | | |
| sw_9 | .07 | .04 | .05 | .09 | .04 | -.02 | .08 | .13 | .09 | .08 | – | | | | | | |
| gw_13_u | .00 | .21 | .35 | -.06 | -.21 | .24 | -.04 | -.22 | .22 | .36 | .14 | – | | | | | |
| ew_2 | .16 | -.16 | -.13 | .21 | .18 | -.11 | .22 | .19 | -.08 | -.19 | -.02 | -.27 | – | | | | |
| gw_4_u | .00 | .33 | .12 | .11 | -.01 | .33 | .07 | -.03 | -.09 | .34 | .05 | .10 | .07 | – | | | |
| gw_16 | .08 | -.12 | -.04 | .11 | .23 | -.09 | .10 | .27 | .05 | -.08 | .19 | -.11 | .06 | .06 | – | | |
| rw_29_u | -.11 | .38 | .20 | .00 | -.15 | .42 | -.01 | -.13 | -.05 | .38 | -.03 | .14 | -.14 | .38 | -.07 | – | |
| ew_16_u | -.21 | .20 | .14 | -.12 | -.10 | .18 | **-.36** | -.06 | .05 | .16 | .08 | .11 | -.15 | .02 | .12 | .16 | – |

Die stärkste negative und positive Korrelation ist hervorgehoben.

TABELLE F.5: Itemkorrelationsmatrix der Teilstichprobe S2 ($N$ = 458) – Teil 1

| Items | gw_17 | rw_33_u | gw_3_u | ew_5_u | rw_11 | sw_2_u | qw_21_u | ew_13 | sw_11_u | rw_18_u | rw_27 | ew_8 | rw_8_u | sw_17_u | sw_22 | sw_15_u | rw_12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| gw_17 | – | | | | | | | | | | | | | | | | |
| rw_33_u | .03 | – | | | | | | | | | | | | | | | |
| gw_3_u | .00 | .24 | – | | | | | | | | | | | | | | |
| ew_5_u | -.01 | .17 | .10 | – | | | | | | | | | | | | | |
| rw_11 | .10 | -.21 | -.03 | .02 | – | | | | | | | | | | | | |
| sw_2_u | -.11 | .29 | .23 | .14 | -.08 | – | | | | | | | | | | | |
| qw_21_u | .02 | .37 | .21 | .23 | -.18 | .10 | – | | | | | | | | | | |
| ew_13 | .09 | .03 | .00 | -.25 | .16 | .05 | -.20 | – | | | | | | | | | |
| sw_11_u | .00 | -.02 | .06 | .02 | -.04 | .20 | .02 | .03 | – | | | | | | | | |
| rw_18_u | .04 | .21 | .16 | .20 | -.10 | .08 | .36 | -.10 | -.01 | – | | | | | | | |
| rw_27 | .25 | -.22 | -.09 | -.08 | .19 | -.14 | -.12 | .10 | -.02 | -.03 | – | | | | | | |
| ew_8 | .04 | -.05 | -.02 | -.15 | .12 | .01 | -.07 | .28 | .04 | -.09 | .10 | – | | | | | |
| rw_8_u | .05 | .21 | .14 | .09 | -.01 | -.06 | .25 | .01 | .00 | .24 | .01 | .03 | – | | | | |
| sw_17_u | -.11 | .29 | .27 | .09 | -.06 | .39 | .18 | -.01 | .17 | .14 | -.13 | -.01 | .04 | – | | | |
| sw_22 | .12 | -.03 | -.04 | -.03 | .19 | -.05 | -.03 | .17 | .04 | -.07 | .27 | .03 | -.01 | .05 | – | | |
| sw_15_u | -.09 | .16 | .16 | .08 | -.04 | .39 | -.02 | .08 | .25 | .02 | -.11 | .06 | -.13 | .29 | -.07 | – | |
| rw_12 | .04 | -.11 | -.02 | -.01 | .26 | -.04 | -.12 | .10 | .01 | -.02 | .25 | .03 | -.05 | .00 | .29 | -.09 | – |
| ew_3 | .02 | -.06 | .00 | -.14 | .17 | .00 | -.09 | .42 | .04 | -.17 | .08 | .47 | .01 | .05 | .13 | .12 | .06 |
| qw_25_u | -.07 | .44 | .17 | .17 | -.21 | .09 | .43 | -.19 | .03 | .33 | -.20 | -.09 | .24 | .23 | -.08 | .03 | -.05 |
| rw_36_u | -.17 | .22 | .26 | .18 | -.05 | .33 | .23 | -.12 | .12 | .21 | -.14 | -.01 | .03 | .29 | -.08 | .22 | .02 |
| ew_10 | .12 | -.05 | .00 | -.17 | .21 | -.05 | -.04 | .28 | -.09 | -.04 | .17 | .39 | -.01 | -.01 | .14 | .00 | .20 |
| gw_9 | .12 | -.18 | -.08 | -.07 | .19 | -.13 | -.12 | .15 | -.02 | -.06 | .21 | .04 | -.06 | -.15 | .23 | -.08 | .10 |
| qw_26_u | -.02 | .39 | .15 | .18 | -.13 | .05 | .38 | -.13 | -.02 | .37 | -.07 | -.09 | .34 | .16 | -.08 | -.02 | .02 |
| ew_12 | .11 | -.09 | -.01 | -.14 | .11 | .11 | -.19 | .43 | .10 | -.09 | .09 | .28 | -.01 | .06 | .13 | .12 | .05 |
| qw_11 | .21 | -.20 | -.01 | -.02 | .40 | -.16 | -.15 | .12 | .01 | -.02 | .40 | .09 | -.06 | -.10 | .33 | -.12 | .43 |
| sw_3_u | -.09 | -.04 | .08 | .00 | -.03 | .16 | .00 | -.04 | .19 | .01 | -.04 | .07 | -.16 | .17 | .00 | .23 | -.05 |
| qw_13_u | -.10 | .32 | .24 | .26 | -.13 | .10 | .34 | -.11 | -.02 | .30 | -.15 | -.11 | .25 | .19 | -.04 | .02 | .01 |
| sw_9 | .02 | .03 | .03 | .00 | .09 | .10 | .03 | .09 | .14 | .04 | .06 | .09 | -.01 | .11 | .17 | .08 | .13 |
| gw_13_u | -.16 | .26 | .27 | .13 | -.15 | .32 | .14 | -.06 | .11 | .13 | -.18 | .03 | -.06 | .36 | -.19 | .34 | -.08 |
| ew_2 | .11 | -.13 | -.10 | -.07 | .05 | -.05 | -.19 | .20 | .05 | -.05 | .16 | .16 | .01 | -.12 | .10 | -.03 | .02 |
| gw_4_u | .04 | .26 | .07 | .21 | -.08 | .03 | .26 | -.06 | .10 | .26 | .04 | -.06 | .20 | .10 | .07 | .01 | -.01 |
| gw_16 | .10 | -.12 | .02 | .08 | .15 | -.09 | .02 | .11 | -.06 | .02 | .14 | .04 | -.10 | -.08 | .15 | -.09 | .11 |
| rw_29_u | -.05 | .31 | .20 | .14 | -.12 | -.04 | .40 | -.08 | -.07 | .46 | -.07 | -.07 | .30 | .10 | -.09 | -.04 | .01 |
| ew_16_u | -.10 | .12 | .12 | .23 | -.03 | .07 | .24 | -.25 | -.01 | .23 | -.03 | -.21 | .13 | .15 | -.03 | -.03 | .01 |

TABELLE F.6: Itemkorrelationsmatrix der Teilstichprobe S2 ($N$ = 458) – Teil 2

| Items | ew_3 | qw_25_u | rw_36_u | ew_10 | gw_9 | qw_26_u | ew_12 | qw_11 | sw_3_u | qw_13_u | sw_9 | gw_13_u | ew_2 | gw_4_u | gw_16 | rw_29_u |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| gw_17 | | | | | | | | | | | | | | | | |
| rw_33_u | | | | | | | | | | | | | | | | |
| gw_3_u | | | | | | | | | | | | | | | | |
| ew_5_u | | | | | | | | | | | | | | | | |
| rw_11 | | | | | | | | | | | | | | | | |
| sw_2_u | | | | | | | | | | | | | | | | |
| qw_21_u | | | | | | | | | | | | | | | | |
| ew_13 | | | | | | | | | | | | | | | | |
| sw_11_u | | | | | | | | | | | | | | | | |
| rw_18_u | | | | | | | | | | | | | | | | |
| rw_27 | | | | | | | | | | | | | | | | |
| ew_8 | | | | | | | | | | | | | | | | |
| rw_8_u | | | | | | | | | | | | | | | | |
| sw_17_u | | | | | | | | | | | | | | | | |
| sw_22 | | | | | | | | | | | | | | | | |
| sw_15_u | | | | | | | | | | | | | | | | |
| rw_12 | | | | | | | | | | | | | | | | |
| ew_3 | – | | | | | | | | | | | | | | | |
| qw_25_u | -.21 | – | | | | | | | | | | | | | | |
| rw_36_u | -.04 | .19 | – | | | | | | | | | | | | | |
| ew_10 | .32 | -.07 | -.04 | – | | | | | | | | | | | | |
| gw_9 | .12 | -.17 | -.18 | .20 | – | | | | | | | | | | | |
| qw_26_u | -.17 | **.55** | .22 | -.05 | -.11 | – | | | | | | | | | | |
| ew_12 | .33 | -.26 | -.07 | .22 | .15 | -.22 | – | | | | | | | | | |
| qw_11 | .12 | -.18 | -.07 | .26 | .28 | -.09 | .12 | – | | | | | | | | |
| sw_3_u | .05 | -.09 | .15 | -.04 | -.06 | -.18 | .09 | -.07 | – | | | | | | | |
| qw_13_u | -.16 | .40 | .28 | -.12 | -.14 | .48 | -.21 | -.09 | .00 | – | | | | | | |
| sw_9 | .06 | .05 | .03 | .13 | .09 | .04 | .05 | .14 | -.01 | .12 | – | | | | | |
| gw_13_u | .02 | .17 | .34 | -.10 | -.18 | .12 | -.02 | -.15 | .21 | .23 | -.01 | – | | | | |
| ew_2 | .19 | -.19 | -.04 | .18 | .12 | -.17 | .30 | .13 | .00 | -.21 | .01 | -.16 | – | | | |
| gw_4_u | -.17 | .34 | .14 | -.02 | .03 | .27 | -.09 | .01 | -.04 | .25 | .08 | .06 | .09 | – | | |
| gw_16 | .03 | -.06 | -.05 | .17 | .16 | -.06 | .05 | .21 | -.06 | .01 | .25 | -.14 | .10 | .04 | – | |
| rw_29_u | -.15 | .41 | .22 | -.03 | -.04 | .47 | -.14 | -.06 | -.11 | .36 | .05 | .09 | -.15 | .32 | .06 | – |
| ew_16_u | -.21 | .27 | .17 | -.16 | -.04 | .29 | **-.31** | .00 | .01 | .29 | .02 | .15 | -.16 | .16 | -.02 | .26 |

Die stärkste negative und positive Korrelation ist hervorgehoben.

## Residualkovarianzmatrizen Studie 2

TABELLE F.7: Geschätzte Residualkovarianzmatrix des initialen 34-Item vierfaktoriellen Modells der Gesamtstichprobe ($N$ = 915) – Teil 1

| Items | qw_26_u | rw_29_u | rw_33_u | qw_25_u | qw_21_u | qw_13_u | rw_8_u | gw_4_u | rw_18_u | gw_13_u | sw_2_u | sw_15_u | gw_3_u | sw_3_u | sw_11_u |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| qw_26_u | 0.000 | | | | | | | | | | | | | | |
| rw_29_u | 0.363 | -0.001 | | | | | | | | | | | | | |
| rw_33_u | 0.158 | -0.890 | 0.000 | | | | | | | | | | | | |
| qw_25_u | 1.273 | -1.251 | 0.455 | 0.000 | | | | | | | | | | | |
| qw_21_u | -1.570 | 0.122 | 0.919 | -0.127 | 0.002 | | | | | | | | | | |
| qw_13_u | 0.413 | -0.101 | -0.377 | -0.283 | -0.822 | 0.000 | | | | | | | | | |
| rw_8_u | 0.391 | 0.850 | -0.007 | -0.270 | 0.923 | -1.177 | 0.000 | | | | | | | | |
| gw_4_u | -0.861 | 1.349 | -0.360 | 0.174 | 0.326 | 0.408 | 0.793 | -0.004 | | | | | | | |
| rw_18_u | -1.121 | 2.633 | -1.844 | -1.230 | 1.694 | -0.597 | 1.703 | 0.246 | 0.000 | | | | | | |
| gw_13_u | 0.457 | -0.986 | **3.141** | 0.714 | -0.471 | **4.137** | **-4.059** | -0.807 | 0.840 | -0.002 | | | | | |
| sw_2_u | -1.665 | **-4.596** | **3.091** | -1.447 | -1.367 | 0.506 | **-3.252** | **-2.486** | -0.826 | -1.126 | -0.001 | | | | |
| sw_15_u | **-2.364** | **-4.082** | 0.773 | **-3.043** | **-3.286** | -0.872 | **-5.240** | **-2.697** | -1.163 | -0.091 | 1.213 | 0.003 | | | |
| gw_3_u | 1.227 | **2.041** | **4.962** | 1.038 | **3.223** | **3.859** | **2.054** | -0.500 | **2.871** | 0.628 | -0.865 | -1.004 | 0.000 | | |
| sw_3_u | **-4.743** | **-4.071** | **-2.860** | **-4.432** | **-3.134** | -0.707 | **-5.626** | **-3.504** | **-1.987** | 1.370 | -1.022 | **3.085** | -0.612 | -0.002 | |
| sw_11_u | **-2.488** | **-3.394** | -0.835 | -1.316 | -0.761 | -0.223 | -0.790 | 0.387 | -1.087 | -1.684 | 0.651 | **2.664** | -0.658 | **2.512** | 0.000 |
| sw_17_u | 0.377 | **-2.226** | **3.040** | 1.171 | 0.364 | **2.547** | **-2.206** | 0.235 | 1.036 | -0.497 | 1.454 | -0.568 | 0.120 | -0.626 | 0.786 |
| rw_36_u | **3.578** | **2.680** | **3.497** | **2.803** | **2.720** | **6.117** | -0.853 | 1.437 | **3.049** | 1.045 | -0.278 | -1.714 | 0.503 | -0.319 | -1.156 |
| qw_11 | -0.490 | 0.433 | **-3.072** | -1.464 | -0.323 | -1.256 | **2.092** | **1.977** | 1.442 | -1.145 | -0.884 | -0.753 | 0.809 | 0.034 | **3.168** |
| rw_27 | -1.254 | -0.309 | **-4.589** | **-2.450** | -1.449 | -1.329 | **2.859** | **2.724** | 1.553 | **-3.105** | -1.187 | -1.477 | -0.604 | -0.693 | 0.603 |
| gw_9 | **-2.289** | -0.560 | **-2.936** | **-2.113** | -0.090 | -1.541 | **2.094** | 1.698 | -0.098 | **-2.954** | -1.206 | -1.481 | -0.844 | -0.296 | 1.271 |
| sw_9 | 1.467 | 1.230 | **2.342** | **2.452** | 1.569 | **3.973** | 0.907 | **2.490** | **2.210** | **3.025** | **4.371** | **3.068** | **2.457** | 1.340 | **3.313** |
| gw_16 | -0.532 | 1.047 | **-2.766** | -0.869 | 1.223 | 0.381 | 0.222 | **2.176** | 1.100 | -1.410 | -0.326 | -0.038 | 1.212 | 0.842 | 0.645 |
| rw_12 | 1.483 | 0.421 | -1.770 | 0.357 | -0.492 | **2.524** | 1.192 | **2.324** | 1.148 | -0.218 | 1.713 | -0.210 | 0.061 | -0.360 | **2.284** |
| rw_11 | -1.535 | -1.022 | **-3.423** | **-2.625** | -1.570 | -1.132 | 1.733 | -0.499 | -0.463 | -1.712 | -0.111 | -0.483 | 1.036 | 1.111 | 1.064 |
| sw_22 | 0.328 | -0.083 | -0.048 | 0.708 | 0.971 | 0.542 | **2.821** | **3.419** | 0.964 | -1.655 | **2.578** | -0.036 | 0.971 | 1.134 | **2.098** |
| ew_10 | 1.575 | 1.636 | 1.523 | -0.145 | 0.457 | -0.246 | **2.183** | **2.999** | 1.257 | 0.232 | 1.618 | 1.416 | 1.277 | -0.663 | -0.078 |
| gw_17 | 1.386 | 1.480 | **2.527** | 0.466 | **3.021** | -1.293 | **4.249** | **3.855** | **2.959** | **-3.973** | -1.598 | -1.857 | -0.210 | **-3.598** | 0.408 |
| ew_12 | 0.099 | 1.911 | 1.558 | -0.610 | -0.015 | 0.026 | **1.984** | **2.628** | 1.650 | 0.206 | 1.845 | **4.196** | 0.111 | **2.066** | **2.114** |
| ew_13 | 1.680 | **2.469** | **3.498** | 0.669 | -0.647 | 0.702 | **3.651** | **3.021** | 1.303 | **-1.998** | 0.500 | **2.006** | -0.201 | -1.121 | 0.829 |
| ew_2 | -1.241 | -1.813 | -1.384 | **-2.031** | **-2.683** | **-3.124** | **2.212** | **4.144** | -0.688 | **-5.831** | -1.864 | -1.365 | **-3.957** | -0.733 | 0.825 |
| ew_16_u | **3.996** | **3.359** | 0.635 | **3.748** | **2.858** | **3.741** | 1.894 | 0.788 | **2.578** | **3.072** | 1.587 | -1.107 | **2.188** | 0.460 | -0.773 |
| ew_5_u | **3.910** | 1.906 | **3.476** | **3.354** | **3.574** | **5.460** | **2.149** | **3.603** | **3.995** | **4.692** | **3.378** | **2.842** | **4.058** | **2.020** | 1.663 |
| ew_3 | 0.891 | 0.113 | 1.681 | -0.499 | 1.053 | 1.133 | 1.765 | 0.560 | 0.065 | 1.526 | 1.441 | **4.306** | **2.389** | 1.656 | 1.577 |
| ew_8 | 1.426 | 1.322 | **2.368** | 0.745 | 0.264 | 1.420 | 1.367 | 0.983 | 0.090 | 0.475 | 1.343 | **3.394** | 0.594 | **2.366** | 0.803 |

Für die Itempaare mit hervorgehobenen Werten liegen signifikante Residuen (-1.96 $\leq z \leq$ 1.96, $\alpha$ = 5%) bzw. als stark zu klassifizierende Abweichungen (-2.58 $\leq z \leq$ 2.58, $\alpha$ = 1%) vor.

TABELLE F.8: Geschätzte Residualkovarianzmatrix des initialen 34-Item vierfaktoriellen Modells der Gesamtstichprobe ($N$ = 915) – Teil 2

| Items | sw_17_u | rw_36_u | qw_11 | rw_27 | gw_9 | sw_9 | gw_16 | rw_12 | rw_11 | sw_22 | ew_10 | gw_17 | ew_12 | ew_13 | ew_2 | ew_16_u | ew_5_u | ew_3 | ew_8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| qw_26_u | | | | | | | | | | | | | | | | | | | |
| rw_29_u | | | | | | | | | | | | | | | | | | | |
| rw_33_u | | | | | | | | | | | | | | | | | | | |
| qw_25_u | | | | | | | | | | | | | | | | | | | |
| qw_21_u | | | | | | | | | | | | | | | | | | | |
| qw_13_u | | | | | | | | | | | | | | | | | | | |
| rw_8_u | | | | | | | | | | | | | | | | | | | |
| gw_4_u | | | | | | | | | | | | | | | | | | | |
| rw_18_u | | | | | | | | | | | | | | | | | | | |
| gw_13_u | | | | | | | | | | | | | | | | | | | |
| sw_2_u | | | | | | | | | | | | | | | | | | | |
| sw_15_u | | | | | | | | | | | | | | | | | | | |
| gw_3_u | | | | | | | | | | | | | | | | | | | |
| sw_3_u | | | | | | | | | | | | | | | | | | | |
| sw_11_u | | | | | | | | | | | | | | | | | | | |
| sw_17_u | 0.016 | | | | | | | | | | | | | | | | | | |
| rw_36_u | -1.024 | 0.003 | | | | | | | | | | | | | | | | | |
| qw_11 | 1.198 | -0.650 | -0.001 | | | | | | | | | | | | | | | | |
| rw_27 | -0.782 | -1.816 | -0.335 | 0.000 | | | | | | | | | | | | | | | |
| gw_9 | -1.516 | **-2.645** | -0.340 | 0.190 | -0.001 | | | | | | | | | | | | | | |
| sw_9 | **5.067** | **2.269** | -0.073 | -1.339 | -0.536 | 0.004 | | | | | | | | | | | | | |
| gw_16 | 0.338 | 0.170 | 0.122 | -1.727 | 1.108 | **4.259** | 0.003 | | | | | | | | | | | | |
| rw_12 | **2.055** | 0.868 | 1.333 | 0.126 | -1.636 | 0.489 | -0.462 | 0.000 | | | | | | | | | | | |
| rw_11 | 0.780 | -0.239 | 0.906 | -1.073 | 0.040 | -0.389 | 0.578 | 0.301 | 0.000 | | | | | | | | | | |
| sw_22 | **4.221** | 0.206 | -1.236 | 1.260 | 0.516 | **2.872** | 0.482 | 1.217 | -0.487 | 0.000 | | | | | | | | | |
| ew_10 | **2.457** | 0.529 | -0.102 | -0.483 | -1.211 | 0.647 | -0.159 | -0.316 | -0.176 | -1.184 | -0.014 | | | | | | | | |
| gw_17 | -1.713 | **-2.499** | -0.481 | 1.476 | 0.181 | -0.120 | 0.080 | -1.374 | -1.602 | 0.675 | 0.205 | 0.000 | | | | | | | |
| ew_12 | **2.543** | -1.551 | -0.387 | -1.074 | 0.832 | 0.188 | -0.306 | **-2.733** | -1.261 | 0.202 | **3.680** | -0.256 | -0.002 | | | | | | |
| ew_13 | -0.286 | **-2.440** | -0.441 | -0.280 | 0.630 | -1.029 | -1.339 | -0.819 | 0.524 | 1.204 | **4.122** | 0.411 | 0.592 | 0.018 | | | | | |
| ew_2 | **-2.919** | -1.951 | 1.722 | **2.679** | **2.425** | -1.132 | 0.908 | -0.323 | 0.361 | 0.540 | **3.791** | **2.988** | 1.064 | -0.333 | -0.003 | | | | |
| ew_16_u | **2.219** | **3.921** | **2.581** | 1.274 | 0.146 | **2.313** | **2.715** | **2.644** | 1.552 | 1.517 | -1.543 | 0.740 | -1.597 | 0.633 | 0.242 | 0.000 | | | |
| ew_5_u | **3.057** | **5.362** | -0.379 | -0.937 | -1.003 | 1.593 | **3.273** | 0.446 | 1.257 | 0.390 | **-2.539** | 1.005 | 0.714 | -0.413 | -0.121 | **2.429** | 0.006 | | |
| ew_3 | 1.643 | -0.331 | -1.248 | -0.773 | 1.119 | 0.471 | -0.649 | -0.422 | 0.768 | 0.124 | **6.155** | -1.343 | -0.951 | 0.948 | -1.050 | 1.473 | 1.753 | -0.005 | |
| ew_8 | 0.904 | 0.132 | -0.293 | 0.289 | -1.108 | 0.312 | -0.933 | -1.282 | 0.681 | -1.644 | **8.474** | -1.618 | -0.663 | -0.356 | -1.186 | 1.016 | 1.375 | **3.000** | 0.001 |

Für die Itempaare mit hervorgehobenen Werten liegen signifikante Residuen (-1.96 $\leq z \leq$ 1.96, $\alpha$ = 5%) bzw. als stark zu klassifizierende Abweichungen (-2.58 $\leq z \leq$ 2.58, $\alpha$ = 1%) vor.

TABELLE F.9: Geschätzte Residualkovarianzmatrix des finalen vierfaktoriellen Modells der Teilstichprobe 1 ($N$ = 457) – Teil 1

| Items | qw_26_u | rw_29_u | rw_33_u | qw_25_u | qw_21_u | qw_13_u | gw_4_u | rw_18_u | gw_13_u | gw_3_u | sw_17_u | sw_2_u | sw_15_u |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| qw_26_u | 0.00 | | | | | | | | | | | | |
| rw_29_u | 0.26 | 0.00 | | | | | | | | | | | |
| rw_33_u | 0.28 | -0.41 | 0.00 | | | | | | | | | | |
| qw_25_u | 1.11 | -0.80 | -0.02 | 0.00 | | | | | | | | | |
| qw_21_u | -0.96 | -0.07 | 0.86 | -0.04 | 0.00 | | | | | | | | |
| qw_13_u | -0.45 | 0.05 | -0.62 | 0.06 | -0.85 | 0.00 | | | | | | | |
| gw_4_u | -0.14 | 1.18 | -0.96 | -0.08 | 0.49 | 0.74 | 0.00 | | | | | | |
| rw_18_u | -1.24 | 1.76 | -0.61 | -0.72 | **2.05** | -0.40 | 0.36 | 0.00 | | | | | |
| gw_13_u | 0.68 | -0.91 | 1.25 | 0.15 | -1.26 | **3.32** | -0.84 | 0.14 | 0.00 | | | | |
| gw_3_u | 0.51 | 0.72 | **3.84** | -0.06 | **2.28** | **2.50** | -0.64 | **2.09** | 0.34 | 0.00 | | | |
| sw_17_u | -0.12 | **-2.82** | 0.33 | -0.28 | -0.77 | 1.84 | -0.11 | 0.37 | -0.68 | -0.19 | 0.00 | | |
| sw_2_u | -0.99 | **-3.33** | 0.61 | -1.10 | -1.43 | 0.52 | **-2.46** | -0.94 | -0.63 | -0.71 | 1.69 | 0.00 | |
| sw_15_u | -0.50 | **-2.48** | -0.29 | **-2.17** | -1.96 | 0.56 | **-2.30** | -0.24 | -0.28 | -0.08 | 0.13 | 0.79 | 0.00 |
| rw_36_u | **2.70** | 1.05 | 1.86 | **2.29** | 1.09 | **4.70** | 0.13 | 1.65 | 1.02 | -0.24 | -1.00 | -0.64 | -1.21 |
| qw_11 | -0.95 | 0.23 | -1.82 | -0.62 | 0.77 | -1.49 | 1.45 | 0.84 | -1.06 | -0.27 | 1.33 | -0.33 | -0.99 |
| rw_27 | -1.74 | -0.21 | **-2.95** | -1.05 | -1.23 | 0.05 | **2.13** | **2.27** | **-1.99** | 0.12 | 0.19 | -0.04 | -1.11 |
| gw_9 | -1.58 | -0.64 | -1.29 | -0.12 | 1.41 | 0.08 | 1.47 | 0.56 | -1.71 | -0.25 | -0.08 | -0.09 | -1.76 |
| gw_16 | -0.14 | -0.09 | **-2.02** | -0.68 | 0.84 | 0.01 | **2.25** | 0.51 | -0.48 | 0.60 | 0.90 | 0.22 | 0.60 |
| rw_12 | 0.46 | -0.64 | -1.25 | 0.22 | 0.56 | **2.55** | **2.76** | 1.33 | -0.27 | -0.50 | 1.49 | 1.83 | -0.08 |
| rw_11 | -0.69 | 0.03 | -1.19 | -0.66 | 0.72 | 0.24 | 0.55 | 0.61 | -0.99 | 0.68 | 0.92 | 0.08 | -1.45 |
| sw_22 | 0.89 | 0.66 | -0.51 | 1.37 | 0.95 | 0.75 | **3.11** | 1.84 | -0.40 | 0.96 | **3.49** | **2.88** | -0.38 |
| ew_10 | 0.78 | 1.20 | 1.66 | -1.11 | -0.26 | 0.42 | **3.09** | 0.91 | -1.00 | -0.37 | 0.16 | -0.27 | -0.90 |
| ew_3 | -0.22 | -1.15 | -0.45 | -1.25 | -0.42 | 0.74 | 1.10 | 0.02 | 0.56 | **2.73** | -0.05 | 0.58 | **2.74** |
| ew_12 | -0.90 | 0.81 | -0.40 | -0.83 | -0.56 | -0.18 | **2.13** | 0.27 | -0.30 | -0.36 | 1.29 | -0.75 | **2.65** |
| ew_13 | -0.18 | 0.62 | -0.05 | -0.37 | -1.03 | -1.22 | 1.87 | 0.10 | **-2.76** | -0.94 | -1.49 | -1.45 | 0.47 |
| ew_8 | -0.35 | -0.46 | 1.19 | -1.04 | -1.70 | 0.75 | -0.35 | -0.88 | -0.58 | 0.85 | 0.90 | 0.93 | **2.93** |

Für die Itempaare mit hervorgehobenen Werten liegen signifikante Residuen (-1.96 $\leq z \leq$ 1.96, $\alpha$ = 5%) bzw. als stark zu klassifizierende Abweichungen (-2.58 $\leq z \leq$ 2.58, $\alpha$ = 1%) vor.

TABELLE F.10: Geschätzte Residualkovarianzmatrix des finalen vierfaktoriellen Modells der Teilstichprobe 1 ($N$ = 457) – Teil 2

| Items | rw_36_u | qw_11 | rw_27 | gw_9 | gw_16 | rw_12 | rw_11 | sw_22 | ew_10 | ew_3 | ew_12 | ew_13 | ew_8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| qw_26_u | | | | | | | | | | | | | |
| rw_29_u | | | | | | | | | | | | | |
| rw_33_u | | | | | | | | | | | | | |
| qw_25_u | | | | | | | | | | | | | |
| qw_21_u | | | | | | | | | | | | | |
| qw_13_u | | | | | | | | | | | | | |
| gw_4_u | | | | | | | | | | | | | |
| rw_18_u | | | | | | | | | | | | | |
| gw_13_u | | | | | | | | | | | | | |
| gw_3_u | | | | | | | | | | | | | |
| sw_17_u | | | | | | | | | | | | | |
| sw_2_u | | | | | | | | | | | | | |
| sw_15_u | | | | | | | | | | | | | |
| rw_36_u | 0.03 | | | | | | | | | | | | |
| qw_11 | -1.78 | 0.00 | | | | | | | | | | | |
| rw_27 | -1.32 | -0.86 | 0.00 | | | | | | | | | | |
| gw_9 | -1.10 | -0.28 | 0.11 | 0.00 | | | | | | | | | |
| gw_16 | 0.60 | 0.69 | -1.67 | 0.78 | 0.00 | | | | | | | | |
| rw_12 | -0.63 | 0.60 | 0.15 | -0.45 | 0.18 | 0.00 | | | | | | | |
| rw_11 | -0.60 | 0.21 | -0.23 | -0.25 | 0.69 | -0.18 | 0.00 | | | | | | |
| sw_22 | 0.68 | -1.08 | 1.56 | 0.02 | 0.56 | 0.77 | 0.07 | 0.00 | | | | | |
| ew_10 | -1.26 | **2.68** | 1.83 | -0.31 | 0.41 | 1.15 | 0.93 | 0.91 | -0.02 | | | | |
| ew_3 | -0.43 | -1.16 | -0.29 | 0.44 | -0.45 | 0.14 | -0.48 | -0.49 | -0.34 | 0.00 | | | |
| ew_12 | **-1.98** | 0.91 | -0.60 | 0.21 | 0.40 | **-2.27** | -1.66 | 0.35 | -0.59 | 0.22 | -0.01 | | |
| ew_13 | **-2.09** | 0.46 | 0.18 | -0.10 | **-2.71** | -0.85 | -0.04 | 0.80 | -1.16 | 0.59 | 1.52 | 0.03 | |
| ew_8 | -0.58 | -0.07 | -0.51 | -1.22 | -0.92 | -1.28 | -0.25 | -1.16 | 0.95 | -0.17 | -0.55 | -0.09 | 0.00 |

Für die Itempaare mit hervorgehobenen Werten liegen signifikante Residuen (-1.96 $\leq z \leq$ 1.96, $\alpha$ = 5%) bzw. als stark zu klassifizierende Abweichungen (-2.58 $\leq z \leq$ 2.58, $\alpha$ = 1%) vor.

TABELLE F.11: Geschätzte Residualkovarianzmatrix des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe 2 ($N$ = 458) – Teil 1

| Items | qw_26_u | rw_29_u | rw_33_u | qw_25_u | qw_21_u | qw_13_u | gw_4_u | rw_18_u | gw_13_u | gw_3_u | sw_17_u | sw_2_u | sw_15_u |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| qw_26_u | 0.00 | | | | | | | | | | | | |
| rw_29_u | 0.40 | 0.00 | | | | | | | | | | | |
| rw_33_u | -0.19 | -0.83 | 0.00 | | | | | | | | | | |
| qw_25_u | 0.72 | -0.88 | 0.53 | 0.00 | | | | | | | | | |
| qw_21_u | -1.05 | 0.47 | 0.49 | -0.02 | 0.00 | | | | | | | | |
| qw_13_u | 0.97 | -0.20 | -0.23 | -0.62 | -0.32 | 0.00 | | | | | | | |
| gw_4_u | -1.01 | 0.88 | 0.38 | 0.41 | 0.24 | -0.24 | 0.00 | | | | | | |
| rw_18_u | -0.12 | **2.24** | -1.92 | -0.80 | 0.83 | -0.39 | 0.17 | 0.00 | | | | | |
| gw_13_u | -0.92 | -1.18 | **2.47** | 0.08 | -0.05 | 1.68 | -0.93 | 0.44 | 0.00 | | | | |
| gw_3_u | 0.65 | 1.63 | **2.71** | 0.89 | 1.93 | **2.41** | -0.48 | 1.61 | 0.27 | 0.00 | | | |
| sw_17_u | -0.25 | -1.08 | **3.06** | 1.10 | 0.63 | 0.90 | -0.23 | 0.49 | -0.07 | 0.18 | 0.00 | | |
| sw_2_u | **-2.45** | **-4.01** | **2.99** | -1.85 | -1.05 | -0.79 | -1.85 | -0.90 | -1.00 | -0.58 | 0.48 | 0.00 | |
| sw_15_u | **-3.27** | **-3.55** | 0.98 | **-2.48** | **-2.84** | **-2.20** | -1.85 | -1.59 | 0.69 | -1.16 | -0.27 | 1.69 | -0.01 |
| rw_36_u | 1.62 | **2.06** | **2.33** | 1.03 | **2.15** | **3.26** | 1.30 | **2.10** | 0.11 | 0.60 | -0.66 | 0.17 | -0.96 |
| qw_11 | 1.21 | 1.17 | -1.84 | -0.61 | -0.63 | 0.50 | **2.01** | 1.86 | -0.22 | 1.71 | 0.66 | -0.55 | -0.12 |
| rw_27 | 0.72 | 0.38 | **-2.78** | -1.71 | -0.39 | -1.08 | **2.33** | 0.70 | **-2.07** | -0.79 | -0.93 | -1.26 | -1.00 |
| gw_9 | -0.96 | 0.46 | **-2.40** | **-2.27** | -1.22 | -1.68 | 1.41 | -0.27 | **-2.15** | -0.71 | -1.71 | -1.20 | -0.33 |
| gw_16 | -0.28 | 1.89 | -1.63 | -0.22 | 1.05 | 0.87 | 1.09 | 1.31 | -1.48 | 1.21 | -0.42 | -0.54 | -0.76 |
| rw_12 | **2.44** | 1.81 | -0.60 | 0.98 | -0.69 | 1.76 | 1.13 | 0.84 | 0.28 | 0.83 | 1.74 | 1.00 | -0.17 |
| rw_11 | -0.80 | -0.92 | **-3.04** | **-2.37** | **-2.40** | -1.20 | -0.70 | -0.72 | -1.11 | 1.01 | 0.52 | 0.15 | 0.76 |
| sw_22 | 0.03 | -0.43 | 0.75 | 0.09 | 0.66 | 0.47 | **2.08** | -0.04 | -1.88 | 0.47 | **2.53** | 0.78 | 0.12 |
| ew_3 | -0.36 | -0.38 | 1.24 | -1.16 | 0.49 | -0.74 | -1.48 | -1.31 | 0.34 | -0.14 | 0.98 | 0.09 | **2.31** |
| ew_12 | -1.68 | -0.48 | 0.26 | **-2.53** | -1.53 | **-1.97** | -0.20 | -0.06 | -0.72 | -0.36 | 1.00 | 1.91 | **2.21** |
| ew_13 | 0.11 | 0.72 | **2.84** | -1.02 | -1.79 | 0.17 | 0.62 | -0.14 | -1.33 | -0.19 | -0.32 | 0.76 | 1.29 |
| ew_8 | 1.20 | 1.10 | 1.25 | 0.95 | 1.10 | 0.27 | 0.78 | 0.02 | 0.27 | -0.69 | -0.56 | -0.10 | 1.07 |
| ew_10 | 1.80 | 1.41 | 0.85 | 1.24 | 1.27 | -0.47 | 1.36 | 1.11 | **-2.22** | -0.08 | -0.34 | -1.24 | -0.09 |

Für die Itempaare mit hervorgehobenen Werten liegen signifikante Residuen (-1.96 $\leq z \leq$ 1.96, $\alpha$ = 5%) bzw. als stark zu klassifizierende Abweichungen (-2.58 $\leq z \leq$ 2.58, $\alpha$ = 1%) vor.

TABELLE F.12: Geschätzte Residualkovarianzmatrix des finalen 26-Item vierfaktoriellen Modells der Teilstichprobe 2 ($N$ = 458) – Teil 2

| Items | rw_36_u | qw_11 | rw_27 | gw_9 | gw_16 | rw_12 | rw_11 | sw_22 | ew_3 | ew_12 | ew_13 | ew_8 | ew_10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| qw_26_u | | | | | | | | | | | | | |
| rw_29_u | | | | | | | | | | | | | |
| rw_33_u | | | | | | | | | | | | | |
| qw_25_u | | | | | | | | | | | | | |
| qw_21_u | | | | | | | | | | | | | |
| qw_13_u | | | | | | | | | | | | | |
| gw_4_u | | | | | | | | | | | | | |
| rw_18_u | | | | | | | | | | | | | |
| gw_13_u | | | | | | | | | | | | | |
| gw_3_u | | | | | | | | | | | | | |
| sw_17_u | | | | | | | | | | | | | |
| sw_2_u | | | | | | | | | | | | | |
| sw_15_u | | | | | | | | | | | | | |
| rw_36_u | -0.01 | | | | | | | | | | | | |
| qw_11 | 0.96 | 0.00 | | | | | | | | | | | |
| rw_27 | -0.96 | 0.24 | 0.00 | | | | | | | | | | |
| gw_9 | **-2.41** | -0.41 | 0.08 | 0.00 | | | | | | | | | |
| gw_16 | -0.35 | -0.22 | -0.42 | 0.87 | 0.00 | | | | | | | | |
| rw_12 | **2.03** | 0.70 | -0.21 | **-2.09** | -0.70 | 0.00 | | | | | | | |
| rw_11 | 0.49 | 0.40 | -1.58 | 0.07 | 0.25 | 0.03 | 0.00 | | | | | | |
| sw_22 | -0.26 | -0.59 | 0.44 | 0.97 | 0.62 | 0.94 | -0.87 | 0.00 | | | | | |
| ew_3 | -1.09 | -1.36 | -1.18 | 0.70 | -0.62 | -1.22 | 1.00 | 0.36 | 0.00 | | | | |
| ew_12 | -1.42 | -0.60 | -0.09 | 1.53 | -0.20 | -1.09 | 0.26 | 0.62 | -0.95 | 0.00 | | | |
| ew_13 | **-2.44** | -0.78 | -0.07 | 1.34 | 1.08 | -0.07 | 1.03 | 1.31 | 0.24 | 1.78 | 0.00 | | |
| ew_8 | -0.03 | -1.40 | 0.14 | -0.94 | -0.69 | -1.32 | 0.46 | -1.63 | 1.05 | -0.64 | -1.55 | 0.01 | |
| ew_10 | -0.78 | **2.78** | 1.89 | **2.26** | **2.37** | **2.35** | **2.37** | 1.25 | -0.69 | -1.00 | -0.57 | **2.12** | 0 |

Für die Itempaare mit hervorgehobenen Werten liegen signifikante Residuen (-1.96 $\leq z \leq$ 1.96, $\alpha$ = 5%) bzw. als stark zu klassifizierende Abweichungen (-2.58 $\leq z \leq$ 2.58, $\alpha$ = 1%) vor.

TABELLE F.13: Geschätzte Residualkovarianzmatrix des finalen 26-Item vierfaktoriellen Modells der Gesamtstichprobe ($N$ = 915) – Teil 1

| Items | qw_26_u | rw_29_u | rw_33_u | qw_25_u | qw_21_u | qw_13_u | gw_4_u | rw_18_u | gw_13_u | gw_3_u | sw_17_u | sw_2_u | sw_15_u |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| qw_26_u | 0.00 | | | | | | | | | | | | |
| rw_29_u | 0.51 | 0.00 | | | | | | | | | | | |
| rw_33_u | 0.04 | -0.94 | 0.00 | | | | | | | | | | |
| qw_25_u | 1.29 | -1.16 | 0.30 | 0.00 | | | | | | | | | |
| qw_21_u | -1.43 | 0.31 | 0.90 | -0.03 | 0.00 | | | | | | | | |
| qw_13_u | 0.37 | -0.07 | -0.57 | -0.37 | -0.77 | 0.00 | | | | | | | |
| gw_4_u | -0.76 | 1.48 | -0.40 | 0.23 | 0.45 | 0.41 | 0.00 | | | | | | |
| rw_18_u | -0.93 | **2.87** | -1.82 | -1.08 | 1.92 | -0.51 | 0.39 | 0.00 | | | | | |
| gw_13_u | -0.22 | -1.56 | **2.53** | 0.02 | -0.99 | **3.52** | -1.27 | 0.36 | 0.00 | | | | |
| gw_3_u | 0.74 | 1.62 | **4.52** | 0.52 | **2.85** | **3.40** | -0.84 | **2.52** | 0.42 | 0.00 | | | |
| sw_17_u | -0.27 | **-2.76** | **2.43** | 0.52 | -0.12 | 1.94 | -0.23 | 0.61 | -0.51 | 0.01 | 0.01 | | |
| sw_2_u | **-2.30** | **-5.13** | **2.51** | **-2.12** | -1.84 | -0.09 | **-2.94** | -1.27 | -1.08 | -0.94 | 1.65 | 0.00 | |
| sw_15_u | **-2.68** | **-4.34** | 0.44 | **-3.39** | **-3.51** | -1.18 | **-2.95** | -1.37 | 0.29 | -0.84 | -0.07 | 1.79 | 0.00 |
| rw_36_u | **2.95** | **2.13** | **2.91** | **2.16** | **2.25** | **5.54** | 1.01 | **2.59** | 0.80 | 0.25 | -1.16 | -0.37 | -1.49 |
| qw_11 | 0.19 | 1.04 | **-2.44** | -0.75 | 0.24 | -0.62 | **2.49** | **1.98** | -0.71 | 1.14 | 1.59 | -0.50 | -0.62 |
| rw_27 | -0.72 | 0.15 | **-4.09** | -1.92 | -1.02 | -0.84 | **3.09** | 1.94 | **-2.81** | -0.38 | -0.53 | -0.95 | -1.42 |
| gw_9 | -1.80 | -0.15 | **-2.50** | -1.60 | 0.30 | -1.08 | **2.02** | 0.24 | **-2.65** | -0.62 | -1.24 | -0.95 | -1.37 |
| gw_16 | -0.26 | 1.26 | **-2.50** | -0.58 | 1.46 | 0.63 | **2.36** | 1.30 | -1.27 | 1.33 | 0.46 | -0.23 | -0.05 |
| rw_12 | **2.03** | 0.90 | -1.29 | 0.90 | -0.05 | **3.02** | **2.68** | 1.52 | 0.10 | 0.30 | **2.33** | **1.99** | -0.10 |
| rw_11 | -0.99 | -0.56 | **-2.92** | **-2.08** | -1.11 | -0.65 | -0.08 | -0.06 | -1.35 | 1.32 | 1.12 | 0.22 | -0.34 |
| sw_22 | 0.73 | 0.25 | 0.32 | 1.10 | 1.25 | 0.91 | **3.67** | 1.23 | -1.47 | 1.12 | **4.35** | **2.70** | -0.03 |
| ew_3 | -0.36 | -1.05 | 0.71 | -1.72 | 0.01 | 0.03 | -0.36 | -0.93 | 0.68 | 1.85 | 0.74 | 0.50 | **3.59** |
| ew_10 | 1.91 | 1.95 | 1.83 | 0.19 | 0.73 | 0.06 | **3.23** | 1.50 | **-2.29** | -0.32 | -0.18 | -1.06 | -0.69 |
| ew_12 | -1.84 | 0.12 | -0.03 | **-2.52** | -1.59 | -1.61 | 1.28 | 0.14 | -0.70 | -0.48 | 1.62 | 0.86 | **3.45** |
| ew_13 | -0.06 | 0.88 | **2.11** | -1.07 | **-2.08** | -0.76 | 1.77 | -0.04 | **-2.86** | -0.76 | -1.15 | -0.38 | 1.33 |
| ew_8 | 0.61 | 0.55 | 1.72 | -0.04 | -0.40 | 0.72 | 0.41 | -0.56 | -0.21 | 0.14 | 0.20 | 0.59 | **2.82** |

Für die Itempaare mit hervorgehobenen Werten liegen signifikante Residuen (-1.96 $\leq z \leq$ 1.96, $\alpha$ = 5%) bzw. als stark zu klassifizierende Abweichungen (-2.58 $\leq z \leq$ 2.58, $\alpha$ = 1%) vor.

TABELLE F.14: Geschätzte Residualkovarianzmatrix des finalen 26-Item vierfaktoriellen Modells der Gesamtstichprobe ($N$ = 915) – Teil 2

| Items | rw_36_u | qw_11 | rw_27 | gw_9 | gw_16 | rw_12 | rw_11 | sw_22 | ew_3 | ew_10 | ew_12 | ew_13 | ew_8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| qw_26_u | | | | | | | | | | | | | |
| rw_29_u | | | | | | | | | | | | | |
| rw_33_u | | | | | | | | | | | | | |
| qw_25_u | | | | | | | | | | | | | |
| qw_21_u | | | | | | | | | | | | | |
| qw_13_u | | | | | | | | | | | | | |
| gw_4_u | | | | | | | | | | | | | |
| rw_18_u | | | | | | | | | | | | | |
| gw_13_u | | | | | | | | | | | | | |
| gw_3_u | | | | | | | | | | | | | |
| sw_17_u | | | | | | | | | | | | | |
| sw_2_u | | | | | | | | | | | | | |
| sw_15_u | | | | | | | | | | | | | |
| rw_36_u | 0.00 | | | | | | | | | | | | |
| qw_11 | -0.23 | 0.00 | | | | | | | | | | | |
| rw_27 | -1.53 | -0.37 | 0.00 | | | | | | | | | | |
| gw_9 | **-2.36** | -0.55 | 0.14 | 0.00 | | | | | | | | | |
| gw_16 | 0.28 | 0.24 | -1.54 | 1.17 | 0.00 | | | | | | | | |
| rw_12 | 1.18 | 1.12 | 0.07 | -1.81 | -0.40 | 0.00 | | | | | | | |
| rw_11 | 0.10 | 0.52 | -1.27 | -0.23 | 0.55 | -0.01 | 0.00 | | | | | | |
| sw_22 | 0.40 | -1.11 | 1.46 | 0.58 | 0.72 | 1.29 | -0.54 | 0.00 | | | | | |
| ew_3 | -1.02 | -1.83 | -1.17 | 0.75 | -0.83 | -0.84 | 0.30 | -0.14 | 0.00 | | | | |
| ew_10 | -1.48 | **3.80** | **2.71** | 1.36 | 1.90 | **2.49** | **2.33** | 1.50 | -0.74 | -0.02 | | | |
| ew_12 | **-2.29** | 0.18 | -0.55 | 1.21 | 0.05 | **-2.34** | -0.95 | 0.65 | -0.46 | -1.12 | 0.00 | | |
| ew_13 | **-3.12** | -0.25 | -0.07 | 0.75 | -1.17 | -0.69 | 0.60 | 1.42 | 0.61 | -1.37 | **2.39** | 0.01 | |
| ew_8 | -0.43 | -1.15 | -0.29 | -1.66 | -1.26 | -1.89 | 0.07 | **-2.06** | 0.57 | **2.13** | -0.90 | -1.29 | 0.00 |

Für die Itempaare mit hervorgehobenen Werten liegen signifikante Residuen (-1.96 $\leq z \leq$ 1.96, $\alpha$ = 5%) bzw. als stark zu klassifizierende Abweichungen (-2.58 $\leq z \leq$ 2.58, $\alpha$ = 1%) vor.

# G Deskriptive Daten zur Kreuzvalidierung

## Deskriptive Daten der StEB-Items für die Teilstichproben

TABELLE G.1: Deskriptive Statistik der StEB-Items von Teilstichprobe 1 ($N$ = 457)

| Items | *n1* | *M* | *SD* | *Mdn* | min | max | range | skew | kurtosis | *SE* |
|---|---|---|---|---|---|---|---|---|---|---|
| gw_17 | 457 | 4.23 | 0.76 | 4 | 2 | 5 | 3 | -0.83 | 0.41 | .04 |
| rw_33_u | 457 | 3.13 | 0.90 | 3 | 1 | 5 | 4 | -0.23 | -0.19 | .04 |
| gw_3_u | 457 | 3.63 | 1.03 | 4 | 1 | 5 | 4 | -0.47 | -0.52 | .05 |
| ew_5_u | 456 | 3.11 | 0.92 | 3 | 1 | 5 | 4 | -0.09 | -0.33 | .04 |
| rw_11 | 457 | 3.96 | 0.94 | 4 | 1 | 5 | 4 | -0.70 | -0.07 | .04 |
| sw_2_u | 455 | 3.91 | 0.98 | 4 | 1 | 5 | 4 | -0.60 | -0.30 | .05 |
| qw_21_u | 456 | 3.03 | 0.78 | 3 | 1 | 5 | 4 | -0.39 | 0.45 | .04 |
| ew_13 | 456 | 3.85 | 0.80 | 4 | 1 | 5 | 4 | -0.38 | -0.13 | .04 |
| sw_11_u | 457 | 3.63 | 0.91 | 4 | 1 | 5 | 4 | -0.26 | -0.32 | .04 |
| rw_18_u | 457 | 3.11 | 0.94 | 3 | 1 | 5 | 4 | -0.25 | -0.42 | .04 |
| rw_27 | 457 | 4.07 | 0.86 | 4 | 1 | 5 | 4 | -0.70 | 0.06 | .04 |
| ew_8 | 456 | 3.25 | 0.84 | 3 | 1 | 5 | 4 | 0.00 | -0.37 | .04 |
| rw_8_u | 457 | 3.68 | 0.80 | 4 | 1 | 5 | 4 | -0.43 | 0.34 | .04 |
| sw_17_u | 456 | 3.71 | 0.95 | 4 | 1 | 5 | 4 | -0.54 | -0.17 | .04 |
| sw_22 | 457 | 3.95 | 0.79 | 4 | 1 | 5 | 4 | -0.72 | 0.72 | .04 |
| sw_15_u | 455 | 3.99 | 0.96 | 4 | 1 | 5 | 4 | -0.52 | -0.61 | .04 |
| rw_12 | 457 | 3.65 | 0.86 | 4 | 1 | 5 | 4 | -0.41 | -0.09 | .04 |
| ew_3 | 457 | 3.39 | 0.80 | 3 | 1 | 5 | 4 | 0.00 | -0.37 | .04 |
| qw_25_u | 457 | 2.61 | 0.94 | 3 | 1 | 5 | 4 | 0.09 | -0.39 | .04 |
| rw_36_u | 455 | 2.97 | 1.10 | 3 | 1 | 5 | 4 | 0.01 | -0.68 | .05 |
| ew_10 | 454 | 3.57 | 0.86 | 4 | 1 | 5 | 4 | -0.25 | -0.40 | .04 |
| gw_9 | 456 | 3.81 | 0.96 | 4 | 1 | 5 | 4 | -0.64 | -0.05 | .05 |
| qw_26_u | 457 | 2.93 | 1.13 | 3 | 1 | 5 | 4 | -0.07 | -0.82 | .05 |
| ew_12 | 456 | 3.72 | 0.82 | 4 | 1 | 5 | 4 | -0.09 | -0.38 | .04 |
| qw_11 | 457 | 3.82 | 0.96 | 4 | 1 | 5 | 4 | -0.66 | -0.07 | .04 |
| sw_3_u | 456 | 2.15 | 0.93 | 2 | 1 | 5 | 4 | 0.76 | 0.39 | .04 |
| qw_13_u | 457 | 2.68 | 0.95 | 3 | 1 | 5 | 4 | 0.25 | -0.14 | .04 |
| sw_9 | 457 | 3.77 | 0.80 | 4 | 1 | 5 | 4 | -0.41 | 0.07 | .04 |
| gw_13_u | 457 | 3.11 | 1.15 | 3 | 1 | 5 | 4 | -0.08 | -0.85 | .05 |
| ew_2 | 456 | 3.55 | 0.95 | 4 | 1 | 5 | 4 | -0.42 | -0.33 | .04 |
| gw_4_u | 455 | 3.52 | 0.82 | 4 | 1 | 5 | 4 | -0.28 | -0.02 | .04 |
| gw_16 | 456 | 3.29 | 0.94 | 3 | 1 | 5 | 4 | -0.23 | -0.31 | .04 |
| rw_29_u | 456 | 3.08 | 0.99 | 3 | 1 | 5 | 4 | -0.24 | -0.48 | .05 |
| ew_16_u | 456 | 2.45 | 0.81 | 3 | 1 | 5 | 4 | -0.36 | -0.47 | .04 |

TABELLE G.2: Deskriptive Statistik der StEB-Items von Teilstichprobe 2 ($N$ = 458)

| Items | *n2* | *M* | *SD* | *Mdn* | min | max | range | skew | kurtosis | *SE* |
|---|---|---|---|---|---|---|---|---|---|---|
| gw_17 | 457 | 4.26 | 0.78 | 4 | 2 | 5 | 3 | -0.88 | 0.33 | .04 |
| rw_33_u | 458 | 3.26 | 0.92 | 3 | 1 | 5 | 4 | -0.21 | -0.14 | .04 |
| gw_3_u | 458 | 3.62 | 1.02 | 4 | 1 | 5 | 4 | -0.43 | -0.41 | .05 |
| ew_5_u | 458 | 3.16 | 0.92 | 3 | 1 | 5 | 4 | -0.18 | -0.35 | .04 |
| rw_11 | 458 | 3.95 | 0.94 | 4 | 1 | 5 | 4 | -0.87 | 0.43 | .04 |
| sw_2_u | 458 | 3.96 | 0.89 | 4 | 1 | 5 | 4 | -0.64 | -0.02 | .04 |
| qw_21_u | 458 | 3.09 | 0.84 | 3 | 1 | 5 | 4 | -0.34 | 0.46 | .04 |
| ew_13 | 458 | 3.91 | 0.82 | 4 | 1 | 5 | 4 | -0.40 | -0.23 | .04 |
| sw_11_u | 458 | 3.61 | 1.00 | 4 | 1 | 5 | 4 | -0.29 | -0.57 | .05 |
| rw_18_u | 458 | 3.17 | 1.01 | 3 | 1 | 5 | 4 | -0.39 | -0.45 | .05 |
| rw_27 | 458 | 4.01 | 0.90 | 4 | 1 | 5 | 4 | -0.72 | -0.02 | .04 |
| ew_8 | 457 | 3.18 | 0.88 | 3 | 1 | 5 | 4 | 0.04 | -0.50 | .04 |
| rw_8_u | 458 | 3.68 | 0.79 | 4 | 1 | 5 | 4 | -0.36 | 0.03 | .04 |
| sw_17_u | 457 | 3.78 | 0.99 | 4 | 1 | 5 | 4 | -0.63 | -0.11 | .05 |
| sw_22 | 458 | 3.95 | 0.80 | 4 | 1 | 5 | 4 | -0.48 | -0.01 | .04 |
| sw_15_u | 457 | 4.02 | 0.98 | 4 | 1 | 5 | 4 | -0.68 | -0.36 | .05 |
| rw_12 | 458 | 3.64 | 0.91 | 4 | 1 | 5 | 4 | -0.34 | -0.31 | .04 |
| ew_3 | 457 | 3.42 | 0.83 | 3 | 1 | 5 | 4 | 0.04 | -0.24 | .04 |
| qw_25_u | 458 | 2.64 | 0.97 | 3 | 1 | 5 | 4 | 0.05 | -0.44 | .05 |
| rw_36_u | 457 | 3.08 | 1.16 | 3 | 1 | 5 | 4 | -0.08 | -0.79 | .05 |
| ew_10 | 458 | 3.51 | 0.86 | 4 | 1 | 5 | 4 | -0.19 | -0.46 | .04 |
| gw_9 | 458 | 3.87 | 0.93 | 4 | 1 | 5 | 4 | -0.53 | -0.48 | .04 |
| qw_26_u | 458 | 2.98 | 1.06 | 3 | 1 | 5 | 4 | -0.04 | -0.74 | .05 |
| ew_12 | 458 | 3.74 | 0.88 | 4 | 1 | 5 | 4 | -0.20 | -0.44 | .04 |
| qw_11 | 457 | 3.88 | 0.92 | 4 | 1 | 5 | 4 | -0.62 | -0.05 | .04 |
| sw_3_u | 457 | 2.19 | 0.92 | 2 | 1 | 5 | 4 | 0.82 | 0.72 | .04 |
| qw_13_u | 458 | 2.77 | 0.96 | 3 | 1 | 5 | 4 | -0.03 | -0.62 | .04 |
| sw_9 | 456 | 3.77 | 0.84 | 4 | 1 | 5 | 4 | -0.55 | 0.27 | .04 |
| gw_13_u | 457 | 3.16 | 1.12 | 3 | 1 | 5 | 4 | 0.04 | -0.78 | .05 |
| ew_2 | 458 | 3.52 | 0.97 | 4 | 1 | 5 | 4 | -0.24 | -0.50 | .05 |
| gw_4_u | 458 | 3.55 | 0.86 | 4 | 1 | 5 | 4 | -0.38 | 0.00 | .04 |
| gw_16 | 458 | 3.33 | 0.89 | 3 | 1 | 5 | 4 | -0.08 | -0.39 | .04 |
| rw_29_u | 458 | 3.07 | 1.03 | 3 | 1 | 5 | 4 | -0.25 | -0.58 | .05 |
| ew_16_u | 458 | 2.41 | 0.82 | 3 | 1 | 5 | 4 | -0.20 | -0.55 | .04 |

# H Modell-Fit der Messmodellprüfungen

TABELLE H.1: Modell-Fit der Messmodellprüfung für die einfaktoriellen Modelle der Gesamtstichprobe ($N = 915$)

| Messmodelle ($N = 915$) | $\chi^2$ | df | $\chi^2$/df | $p$ | CFI | RMSEA | CI (90%) | CFit | SRMR |
|---|---|---|---|---|---|---|---|---|---|
| Absolutheit | | | | | | | | | |
| paralleles | 1028.434 | 41 | 25.08 | .000 | .42 | .162 | .154; .171 | .000 | .176 |
| essentiell paralleles | 240.109 | 34 | 7.06 | .000 | .88 | .081 | .072; .091 | .000 | .110 |
| $\tau$-äuquivalentes | 898.980 | 34 | 26.44 | .000 | .47 | .167 | .158; .176 | .000 | .173 |
| essentiell $\tau$-äuquivalentes | 165.731 | 27 | 6.14 | .000 | **.91** | .075 | .065; .085 | .000 | .083 |
| **$\tau$-kongenerisches** | 70.746 | 20 | **3.54** | .000 | **.97** | **.053** | .041; .065 | **.341*** | **.029** |
| Einfachheit | | | | | | | | | |
| paralleles | 1073.756 | 24 | 44.74 | .000 | .00 | .219 | .208; .230 | .000 | .209 |
| essentiell paralleles | 174.969 | 19 | 9.21 | .000 | .81 | .095 | .082; .108 | .000 | .091 |
| $\tau$-äuquivalentes | 826.806 | 19 | 43.52 | .000 | .00 | .216 | .204; .228 | .000 | .199 |
| essentiell $\tau$-äuquivalentes | 67.755 | 14 | 4.84 | .000 | **.93** | .065 | .051; .080 | .044 | **.066** |
| **$\tau$-kongenerisches** | 26.282 | 9 | **2.92** | .002 | **.98** | **.046** | .027; .065 | **.607*** | **.024** |
| Multimodalität | | | | | | | | | |
| paralleles | 619.920 | 32 | 19.37 | .00 | .20 | .142 | .133; .151 | .000 | .147 |
| essentiell paralleles | 138.651 | 26 | 5.33 | .000 | .84 | .069 | .058; .080 | .002 | .081 |
| $\tau$-äuquivalentes | 491.506 | 26 | 18.9 | .000 | .35 | .140 | .130; .150 | .000 | .155 |
| essentiell $\tau$-äuquivalentes | 88.656 | 20 | 4.43 | .000 | **.90** | .061 | .050; .074 | .057 | **.065** |
| **$\tau$-kongenerisches** | 29.799 | 14 | **2.13** | .008 | **.98** | **.035** | **.019; .051** | **.935*** | **.024** |
| Entwicklung | | | | | | | | | |
| paralleles | 542.498 | 17 | 31.91 | .000 | .23 | .184 | .171; .197 | .000 | .155 |
| essentiell paralleles | 62.278 | 13 | 4.79 | .000 | **.93** | .064 | .049; .081 | .061 | **.048** |
| $\tau$-äuquivalentes | 493.063 | 13 | 37.93 | .000 | .26 | .201 | .186; .216 | .000 | .158 |
| **essentiell $\tau$-äuquivalentes** | 43.879 | 9 | 4.88 | .000 | **.95** | .065 | .047; .084 | .079 | **.042** |
| $\tau$-kongenerisches | 35.625 | 5 | 7.13 | .000 | **.95** | .082 | .059; .106 | .012 | **.034** |

*RMSEA p-close-fit: $.10 \leq p \leq 1.00$ (RMSEA < .05)
hervorgehobene Werte befinden sich innerhalb der Cut-off Werte von Hu und Bentler (1999)

TABELLE H.2: Modell-Fit der Messmodellprüfung der einfaktoriellen Modelle von Teilstichprobe 1 ($N$ = 457)

| Messmodelle S1 ($N$ = 457) | $\chi^2$ | df | $\chi^2$/df | $p$ | CFI | RMSEA | CI (90%) | CFit | SRMR |
|---|---|---|---|---|---|---|---|---|---|
| Absolutheit | | | | | | | | | |
| paralleles | 546.899 | 41 | 13.34 | .000 | .39 | .164 | .153; .176 | .000 | .188 |
| essentiell paralleles | 142.333 | 34 | 4.19 | .000 | .87 | .083 | .070; .097 | .000 | .122 |
| $\tau$-äuquivalentes | 471.394 | 34 | 13.86 | .000 | .45 | .168 | .155; .181 | .000 | .174 |
| essentiell $\tau$-äuquivalentes | 86.253 | 27 | 3.19 | .000 | **.93** | .069 | .054; .085 | .021 | .081 |
| **$\tau$-kongenerisches** | 42.245 | 20 | **2.11** | .003 | **.97** | **.049** | .029; .069 | **.493*** | **.031** |
| Einfachheit | | | | | | | | | |
| paralleles | 598.699 | 24 | 24.95 | .000 | .00 | .229 | .213; .245 | .000 | .216 |
| essentiell paralleles | 115.89 | 19 | 6.10 | .000 | .78 | .106 | .088; .124 | **.124*** | .096 |
| $\tau$-äuquivalentes | 459.312 | 19 | 24.17 | .000 | .00 | .225 | .208; .243 | .000 | .218 |
| essentiell $\tau$-äuquivalentes | 56.789 | 14 | 4.06 | .000 | **.90** | .082 | .061; .104 | .007 | .084 |
| **$\tau$-kongenerisches** | 17.389 | 9 | **1.93** | .043 | **.98** | **.045** | .011; .075 | **.560*** | **.028** |
| Multimodalität | | | | | | | | | |
| paralleles | 327.362 | 32 | 10.23 | .000 | .23 | .142 | .129; .155 | .000 | .153 |
| essentiell paralleles | 68.426 | 26 | **2.63** | .000 | .89 | **.060** | .044; .076 | **.148*** | .085 |
| $\tau$-äuquivalentes | 254.97 | 26 | 9.81 | .000 | .39 | .139 | .125; .153 | .000 | .153 |
| essentiell $\tau$-äuquivalentes | 37.979 | 20 | **1.90** | .009 | **.95** | **.044** | .024; .064 | **.658*** | **.060** |
| **$\tau$-kongenerisches** | 18.787 | 14 | **1.34** | **.173** | **.99** | **.027** | **.000; .054** | **.915*** | **.027** |
| Entwicklung | | | | | | | | | |
| paralleles | 230.702 | 17 | 13.57 | .000 | .26 | .166 | .147; .185 | .000 | .146 |
| **essentiell paralleles** | 21.983 | 13 | **1.69** | **.056** | **.97** | **.039** | .000; .066 | **.722*** | **.045** |
| $\tau$-äuquivalentes | 210.684 | 13 | 16.21 | .000 | .28 | .182 | .162; .204 | .000 | .149 |
| **essentiell $\tau$-äuquivalentes** | 15.092 | 9 | **1.68** | **.088** | **.98** | **.038** | .000; .070 | **.690*** | **.038** |
| **$\tau$-kongenerisches** | 10.576 | 5 | **2.12** | **.060** | **.98** | **.049** | .002; .088 | **.452*** | **.027** |

*RMSEA p-close-fit: $.10 \leq p \leq 1.00$ (RMSEA < .05)
hervorgehobene Werte befinden sich innerhalb der Cut-off Werte von Hu und Bentler (1999)

TABELLE H.3: Modell-Fit der Messmodellprüfung der einfaktoriellen Modelle von Teilstichprobe 2 ($N$ = 458)

| Messmodelle S2 ($N$ = 458) | $\chi^2$ | df | $\chi^2$/df | $p$ | CFI | RMSEA | CI (90%) | CFit | SRMR |
|---|---|---|---|---|---|---|---|---|---|
| Absolutheit | | | | | | | | | |
| paralleles | 522.114 | 41 | 12.73 | .000 | .45 | .160 | .148; .172 | .000 | .171 |
| essentiell paralleles | 133.441 | 34 | 3.92 | .000 | .88 | .080 | .067; .094 | .000 | .106 |
| $\tau$-äuquivalentes | 455.190 | 34 | 13.39 | .000 | .49 | .164 | .152; .177 | .000 | .177 |
| essentiell $\tau$-äuquivalentes | 99.890 | 27 | 3.70 | .000 | **.91** | .077 | .062; .092 | .002 | .093 |
| **$\tau$-kongenerisches** | 44.092 | 20 | **2.20** | .001 | **.97** | **.051** | .032; .070 | **.427*** | **.032** |
| Einfachheit | | | | | | | | | |
| paralleles | 508.656 | 24 | 21.19 | .000 | .00 | .210 | .195; .226 | .000 | .211 |
| essentiell paralleles | 85.059 | 19 | 4.48 | .000 | .83 | .087 | .069; .106 | .001 | .099 |
| $\tau$-äuquivalentes | 392.894 | 19 | 20.68 | .000 | .00 | .207 | .191; .224 | .000 | .187 |
| essentiell $\tau$-äuquivalentes | 28.820 | 14 | **2.06** | .000 | **.96** | **.048** | .024; .072 | .518 | **.055** |
| **$\tau$-kongenerisches** | 16.471 | 9 | **1.83** | **.058** | **.98** | **.043** | .003; .073 | **.616*** | **.027** |
| Multimodalität | | | | | | | | | |
| paralleles | 338.860 | 32 | 10.59 | .000 | .17 | .145 | .131; .158 | .000 | .149 |
| essentiell paralleles | 107.650 | 26 | 4.41 | .000 | .78 | .083 | .068; .099 | .000 | .090 |
| $\tau$-äuquivalentes | 273.696 | 26 | 10.53 | .000 | .30 | .144 | .130; .159 | .000 | .164 |
| essentiell $\tau$-äuquivalentes | 78.216 | 20 | 3.91 | .000 | .84 | .080 | .063; .098 | .003 | .083 |
| **$\tau$-kongenerisches** | 21.528 | 14 | **1.54** | **.089** | **.98** | **.034** | **.000; .060** | **.826*** | **.030** |
| Entwicklung | | | | | | | | | |
| paralleles | 335.421 | 17 | 19.73 | .000 | .20 | .202 | .184; .221 | .000 | .169 |
| essentiell paralleles | 55.795 | 13 | 4.29 | .000 | .89 | .085 | .063; .108 | .006 | **.060** |
| $\tau$-äuquivalentes | 304.749 | 13 | 23.44 | .000 | .23 | .221 | .201; .243 | .000 | .171 |
| **essentiell $\tau$-äuquivalentes** | 41.457 | 9 | 4.61 | .000 | **.91** | .089 | .063; .116 | .008 | **.055** |
| $\tau$-kongenerisches | 35.652 | 5 | 7.13 | .000 | **.92** | .116 | .082; .152 | .001 | **.042** |

*RMSEA p-close-fit: $.10 \leq p \leq 1.00$ (RMSEA < .05)
hervorgehobene Werte befinden sich innerhalb der Cut-off Werte von Hu und Bentler (1999)

# I Testwertvalidierung

## Deskriptiven Datenanalyse der StEB-Items für Prüfung der Grundannahme 4

TABELLE I.1: Deskriptive Analyse der StEB-Items für Prüfung der Grundannahme 4 (bereinigte Stichprobe, $N = 891$)

| Items | *n* | *M* | *SD* | median | min | max | range | skew | kurtosis | *SE* |
|---|---|---|---|---|---|---|---|---|---|---|
| qw_26_u | 891 | 2.97 | 1.09 | 3 | 1 | 5 | 4 | -0.07 | -0.76 | .04 |
| rw_29_u | 890 | 3.07 | 1 | 3 | 1 | 5 | 4 | -0.25 | -0.51 | .03 |
| rw_33_u | 891 | 3.2 | 0.92 | 3 | 1 | 5 | 4 | -0.22 | -0.15 | .03 |
| qw_25_u | 891 | 2.63 | 0.96 | 3 | 1 | 5 | 4 | 0.08 | -0.41 | .03 |
| qw_21_u | 890 | 3.07 | 0.8 | 3 | 1 | 5 | 4 | -0.36 | 0.52 | .03 |
| qw_13_u | 891 | 2.74 | 0.96 | 3 | 1 | 5 | 4 | 0.09 | -0.4 | .03 |
| gw_4_u | 889 | 3.54 | 0.84 | 4 | 1 | 5 | 4 | -0.34 | 0.02 | .03 |
| rw_18_u | 891 | 3.14 | 0.97 | 3 | 1 | 5 | 4 | -0.32 | -0.41 | .03 |
| gw_13_u | 890 | 3.14 | 1.13 | 3 | 1 | 5 | 4 | -0.02 | -0.8 | .04 |
| gw_3_u | 891 | 3.63 | 1.03 | 4 | 1 | 5 | 4 | -0.46 | -0.45 | .04 |
| sw_17_u | 889 | 3.75 | 0.96 | 4 | 1 | 5 | 4 | -0.58 | -0.12 | .03 |
| sw_2_u | 889 | 3.94 | 0.93 | 4 | 1 | 5 | 4 | -0.62 | -0.16 | .03 |
| sw_15_u | 888 | 4.01 | 0.96 | 4 | 1 | 5 | 4 | -0.58 | -0.55 | .03 |
| rw_36_u | 888 | 3.02 | 1.12 | 3 | 1 | 5 | 4 | -0.03 | -0.73 | .04 |
| qw_11 | 890 | 3.84 | 0.94 | 4 | 1 | 5 | 4 | -0.64 | -0.06 | .03 |
| rw_27 | 891 | 4.04 | 0.88 | 4 | 1 | 5 | 4 | -0.69 | -0.05 | .03 |
| gw_9 | 890 | 3.83 | 0.94 | 4 | 1 | 5 | 4 | -0.57 | -0.26 | .03 |
| gw_16 | 890 | 3.3 | 0.91 | 3 | 1 | 5 | 4 | -0.17 | -0.31 | .03 |
| rw_12 | 891 | 3.65 | 0.89 | 4 | 1 | 5 | 4 | -0.37 | -0.17 | .03 |
| rw_11 | 891 | 3.95 | 0.94 | 4 | 1 | 5 | 4 | -0.78 | 0.19 | .03 |
| sw_22 | 891 | 3.96 | 0.79 | 4 | 1 | 5 | 4 | -0.57 | 0.26 | .03 |
| ew_3 | 890 | 3.41 | 0.81 | 3 | 1 | 5 | 4 | 0.01 | -0.25 | .03 |
| ew_10 | 888 | 3.53 | 0.86 | 4 | 1 | 5 | 4 | -0.23 | -0.43 | .03 |
| ew_12 | 890 | 3.73 | 0.83 | 4 | 1 | 5 | 4 | -0.1 | -0.48 | .03 |
| ew_13 | 890 | 3.88 | 0.8 | 4 | 1 | 5 | 4 | -0.36 | -0.27 | .03 |
| ew_8 | 889 | 3.22 | 0.85 | 3 | 1 | 5 | 4 | 0 | -0.4 | .03 |

## Histogramme der Testwerte des StEB

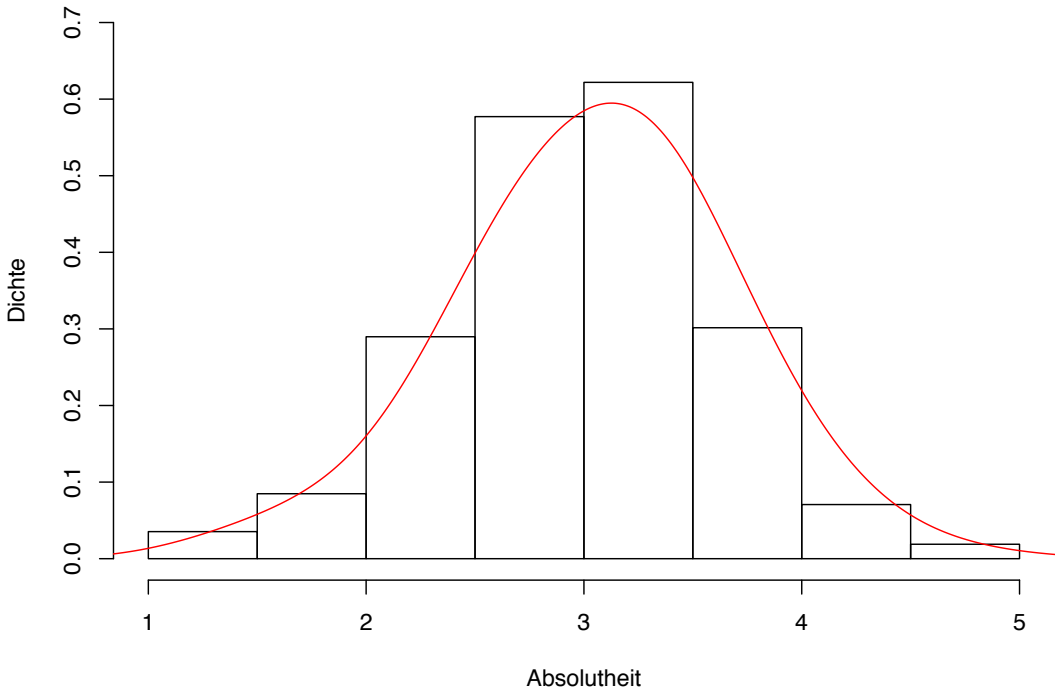


ABBILDUNG I.1: Histogramm der individuellen Skalenmittelwerte zur Absolutheit des Wissens

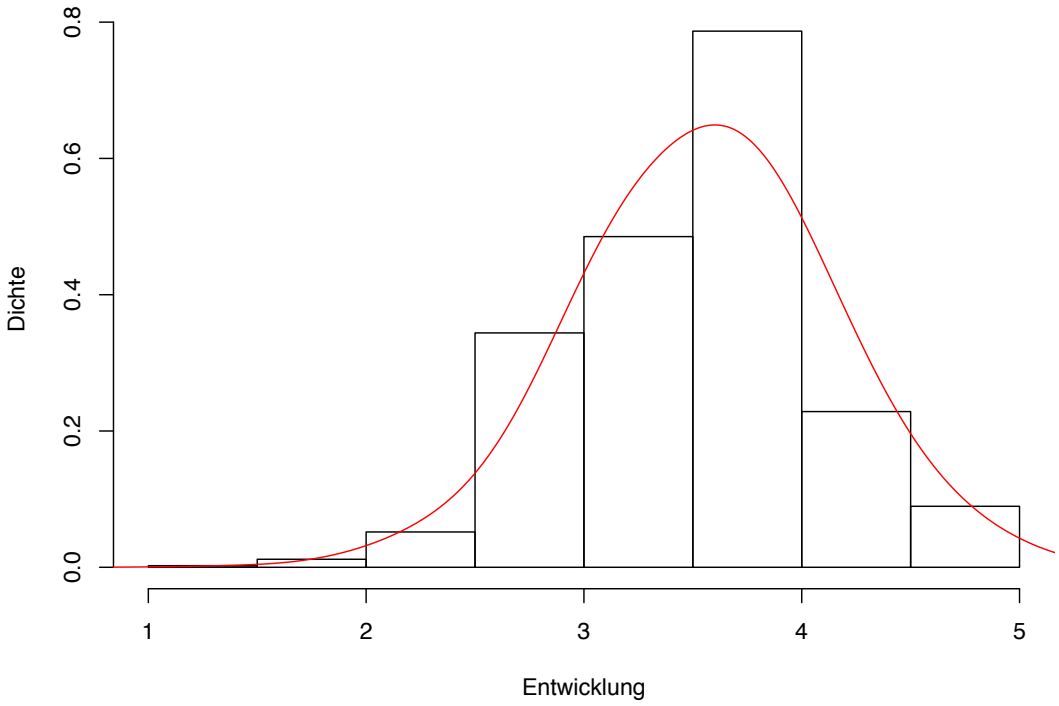


ABBILDUNG I.2: Histogramm der individuellen Skalenmittelwerte zur Entwicklung des Wissens

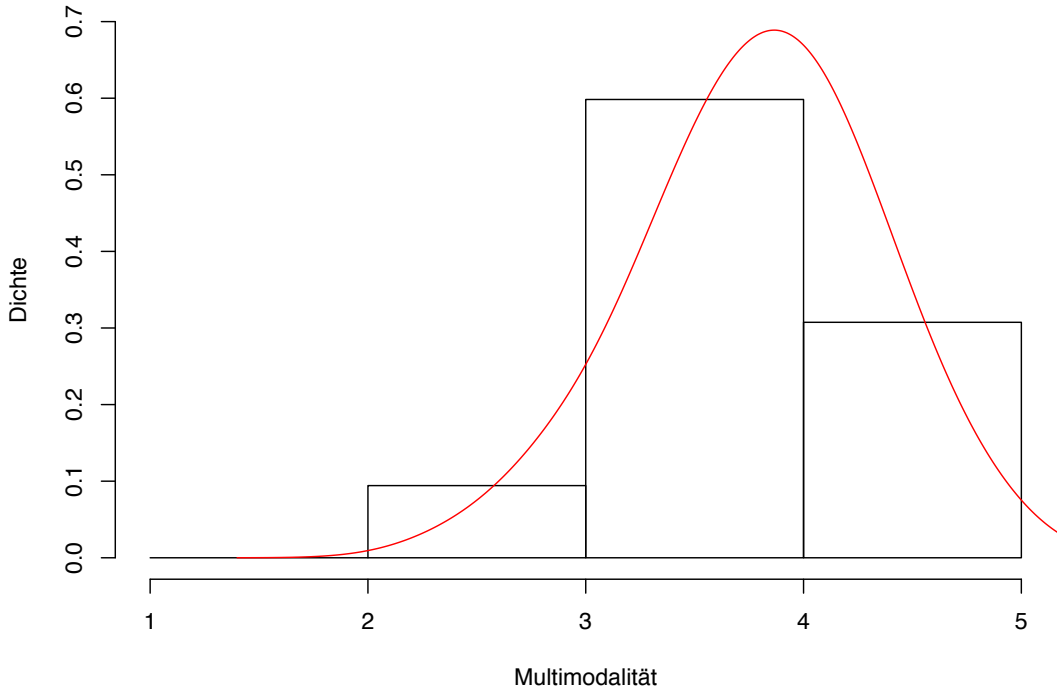


ABBILDUNG I.3: Histogramm der individuellen Skalenmittelwerte zur Multimodalität des Wissenserwerbs

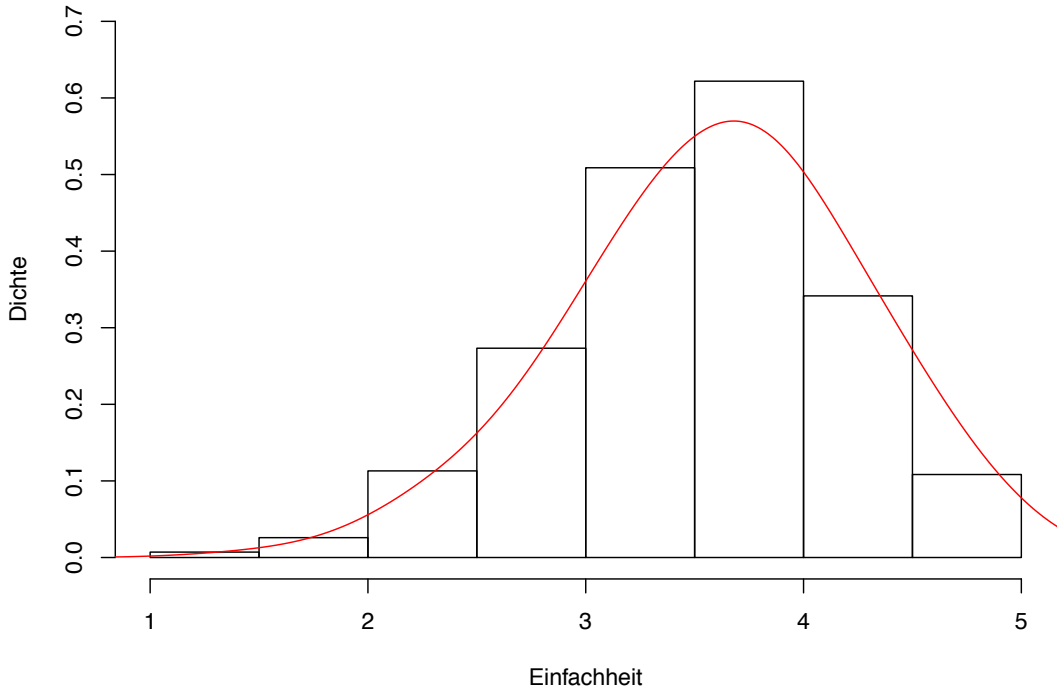


ABBILDUNG I.4: Histogramm der individuellen Skalenmittelwerte zur Einfachheit des Wissens

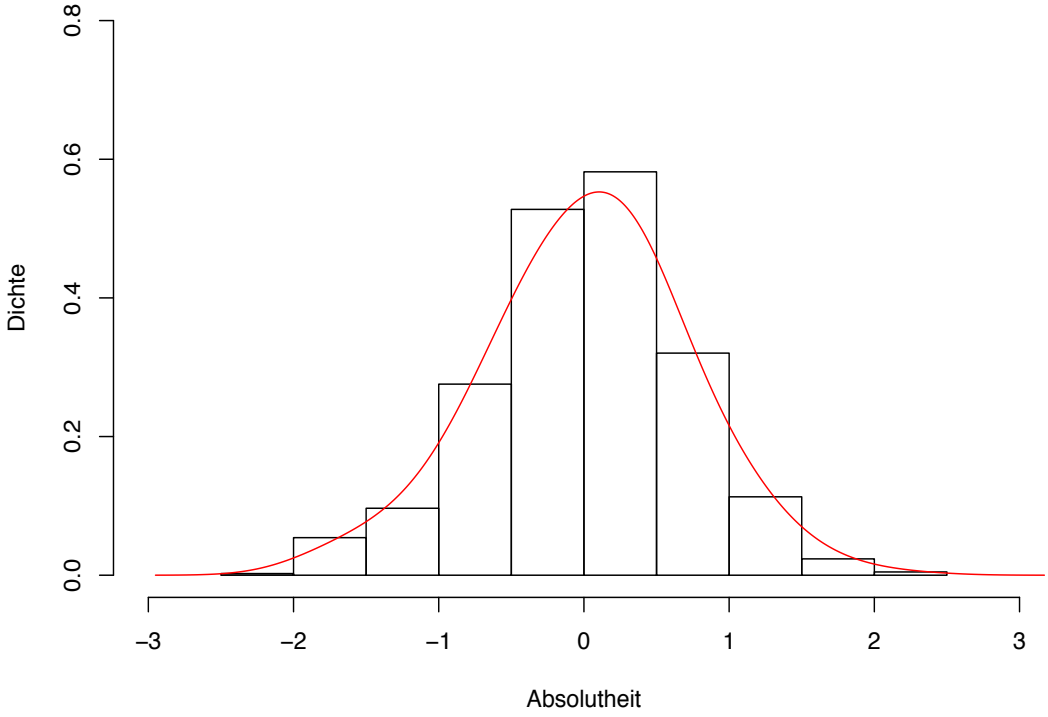


ABBILDUNG I.5: Histogramm der individuellen Faktorwerte zur Absolutheit des Wissens

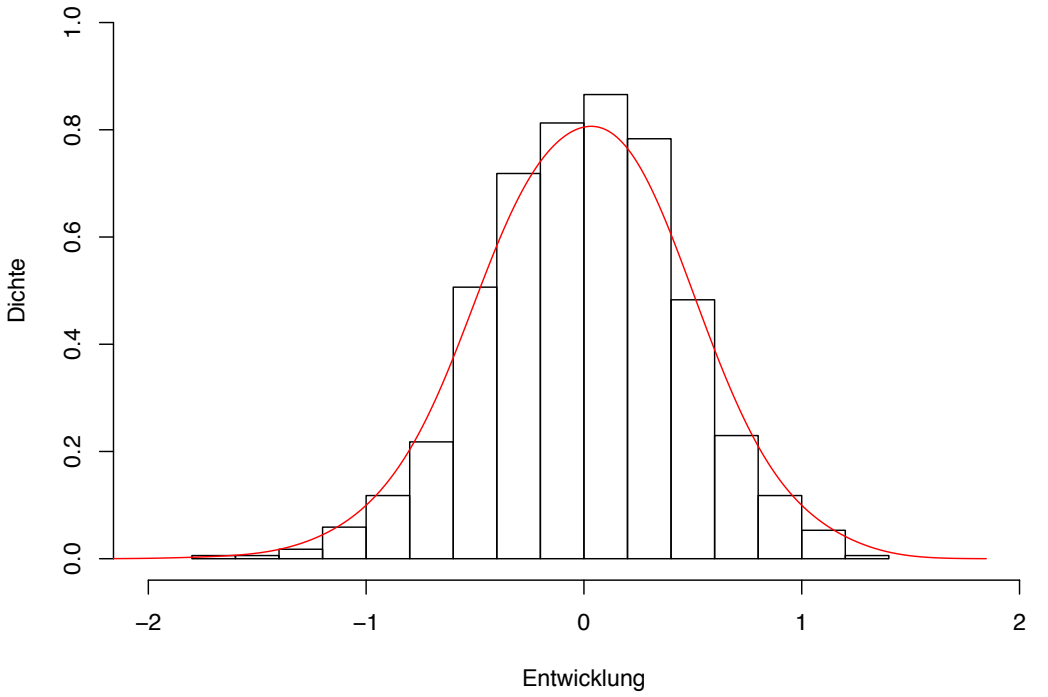


ABBILDUNG I.6: Histogramm der individuellen Faktorwerte zur Entwicklung des Wissens

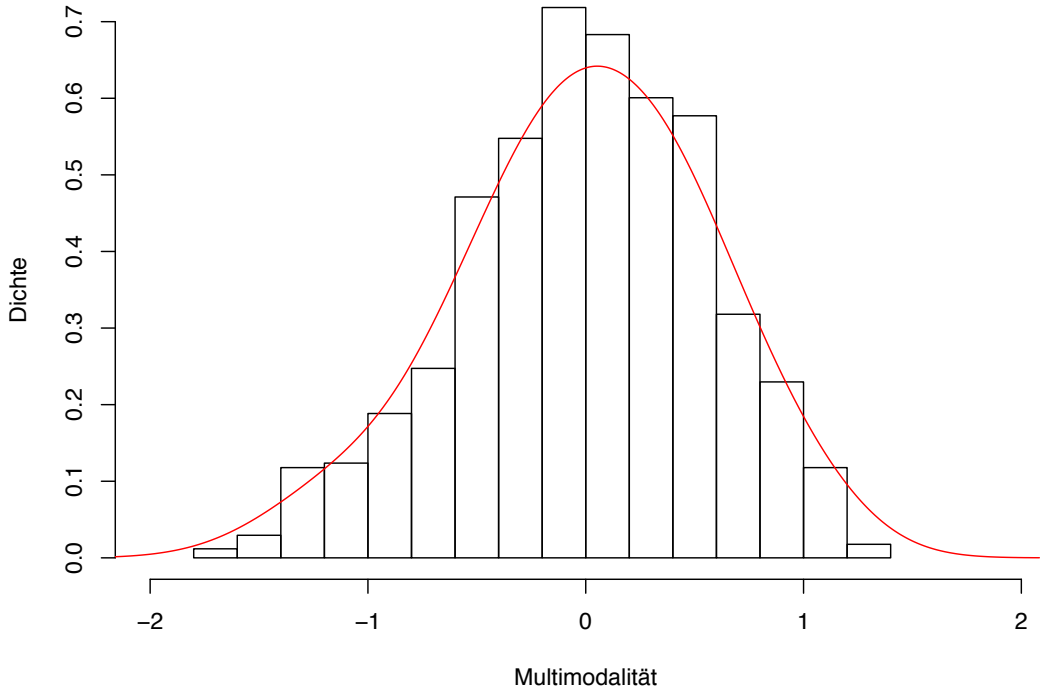


ABBILDUNG I.7: Histogramm der individuellen Faktorwerte zur Multimodalität des Wissenserwerbs

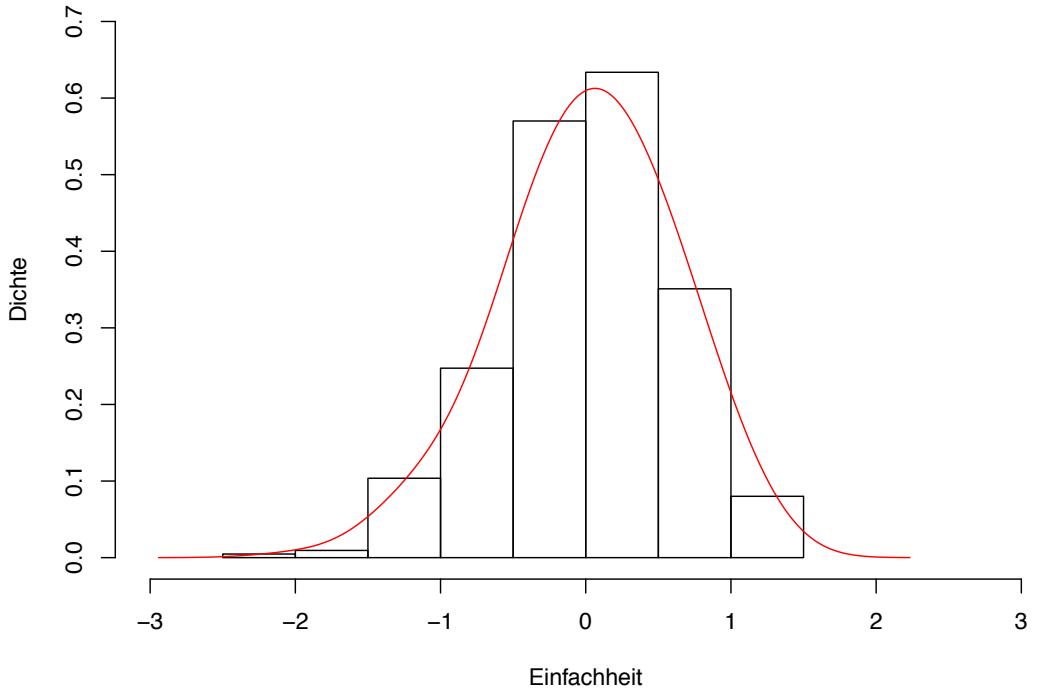


ABBILDUNG I.8: Histogramm der individuellen Faktorwerte zur Einfachheit des Wissens

## Unterschiede in den Überzeugungsausprägungen in Abhängigkeit der Hochschule

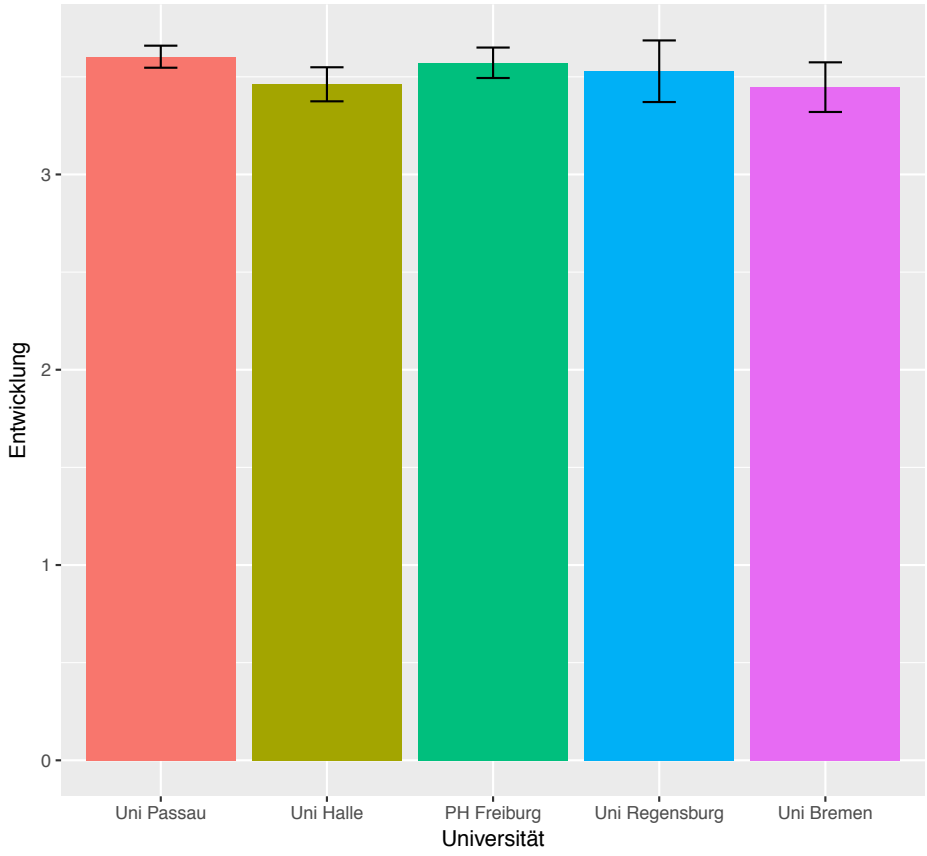


ABBILDUNG I.9: Durchschnittliche Ausprägung in der Überzeugung zur Entwicklung des Wissens nach Hochschule

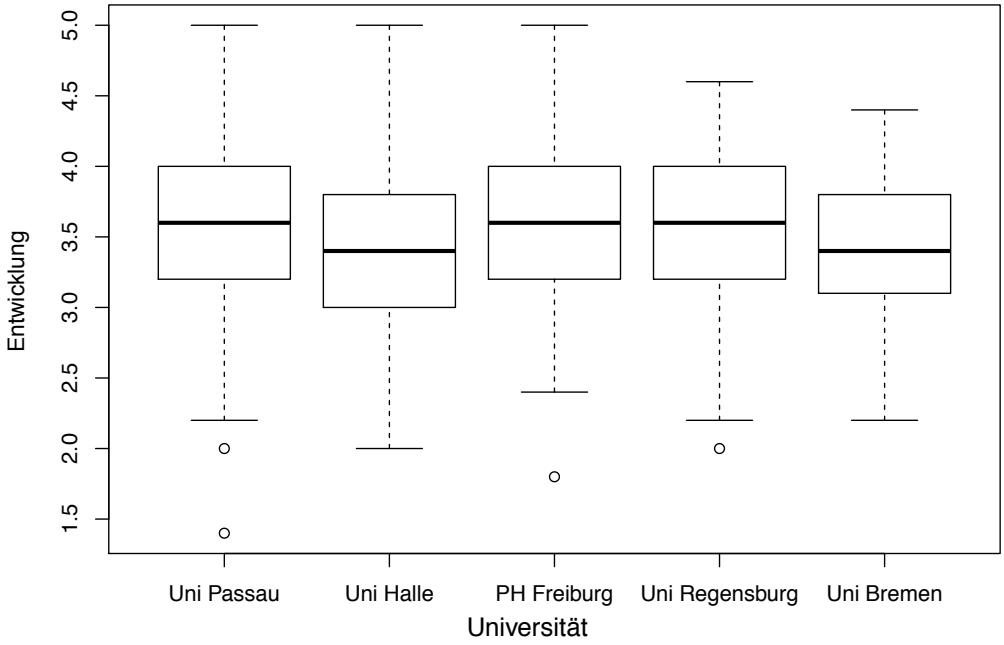


ABBILDUNG I.10: Boxplot der individuellen Skalenmittelwerte zur Entwicklung des Wissens nach Hochschule

## Unterschiede in den Überzeugungsausprägungen in Abhängigkeit der Schulart

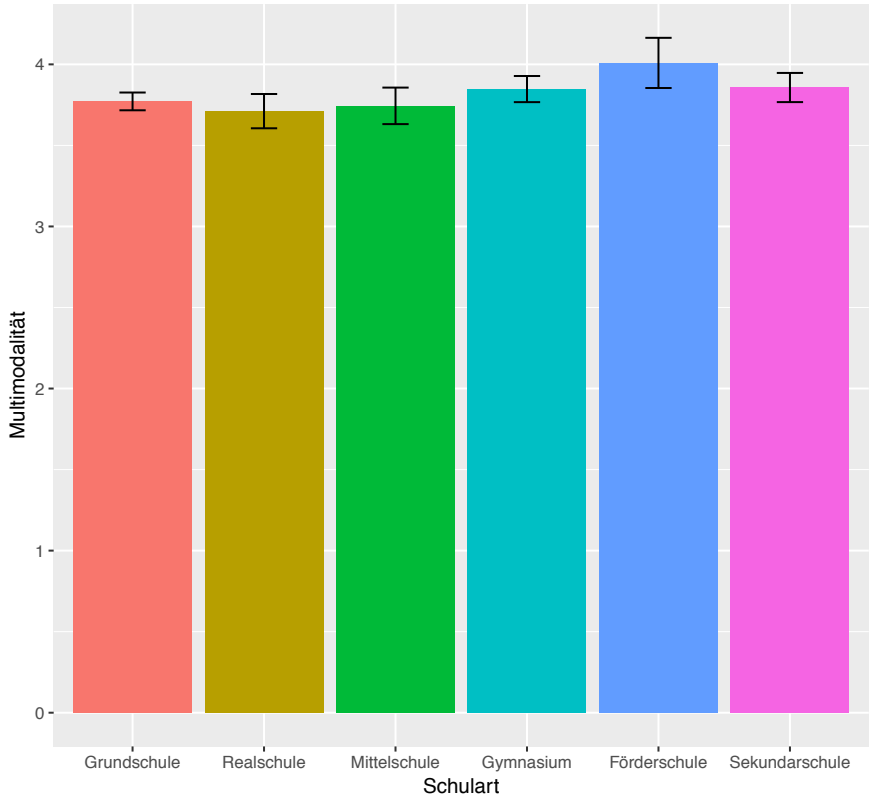


ABBILDUNG I.11: Durchschnittliche Ausprägung in der Überzeugung zur Multimodalität des Wissenserwerbs nach Schulart

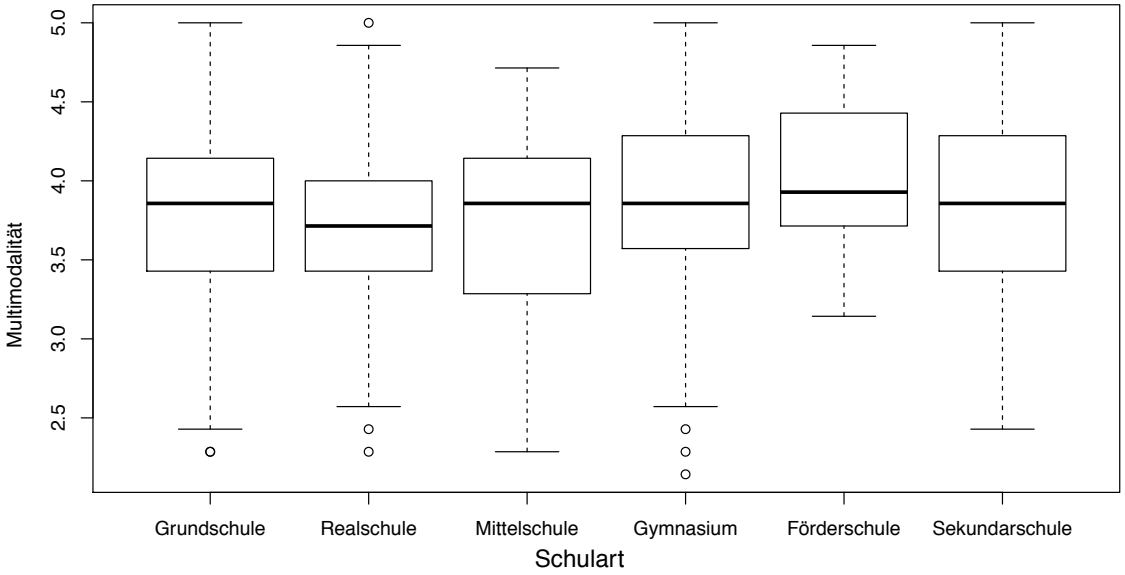


ABBILDUNG I.12: Boxplot der individuellen Skalenmittelwerte zur Multimodalität des Wissenserwerbs nach Schulart

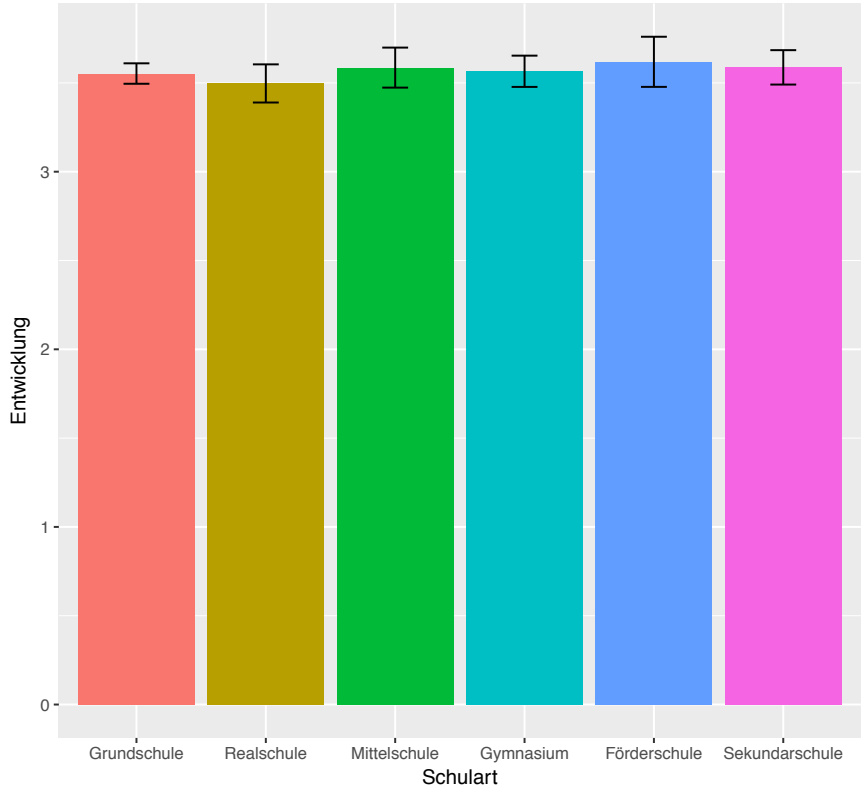


ABBILDUNG I.13: Durchschnittliche Ausprägung in der Überzeugung zur Entwicklung des Wissens nach Schulart

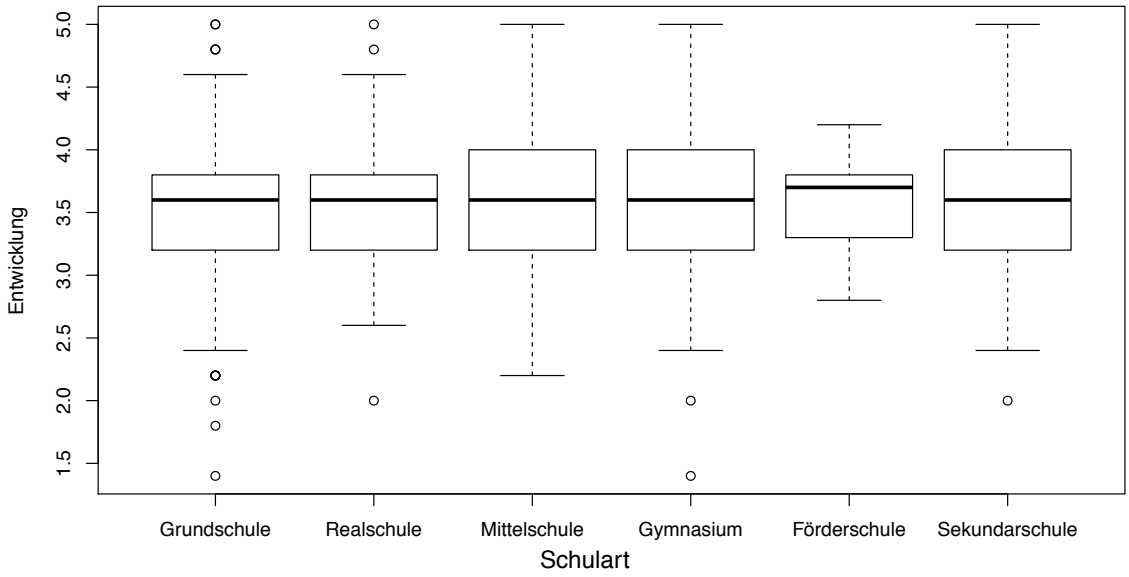


ABBILDUNG I.14: Boxplot der individuellen Skalenmittelwerte zur Entwicklung des Wissens nach Schulart

## Unterschiede in den Überzeugungsausprägungen in Abhängigkeit von der fachlichen Ausrichtung

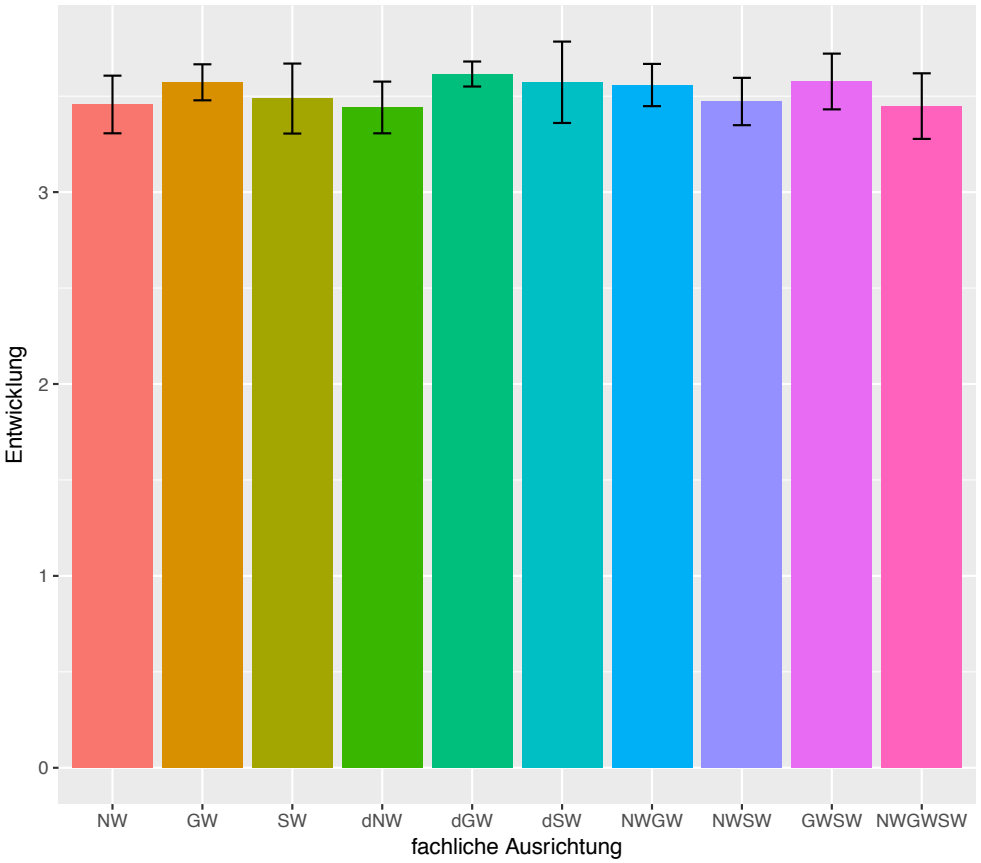


ABBILDUNG I.15: Durchschnittliche Ausprägung in der Überzeugung zur Entwicklung des Wissens nach fachlicher Ausrichtung

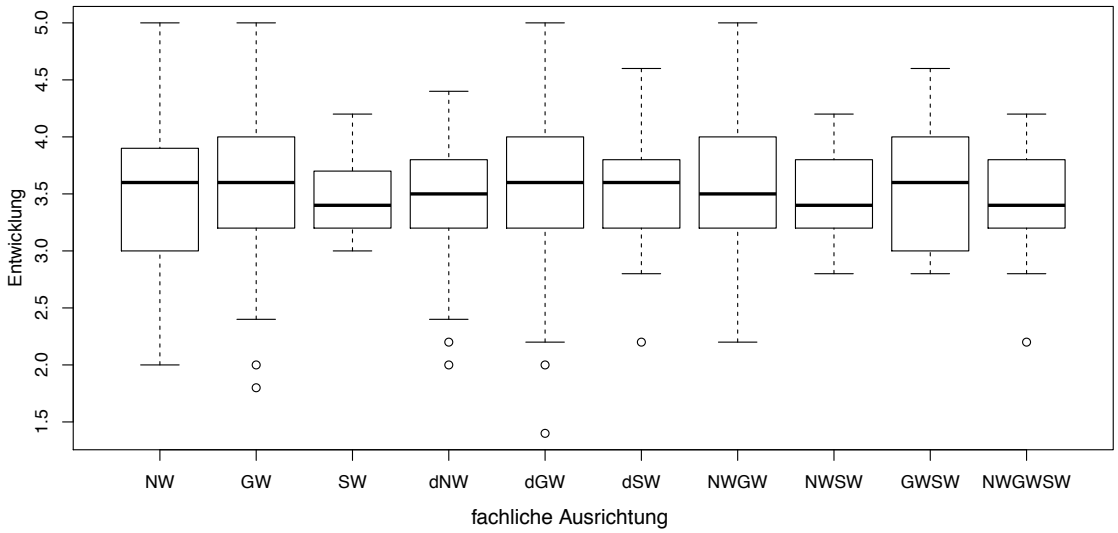


ABBILDUNG I.16: Boxplot der individuellen Skalenmittelwerte zur Entwicklung des Wissens nach fachlicher Ausrichtung

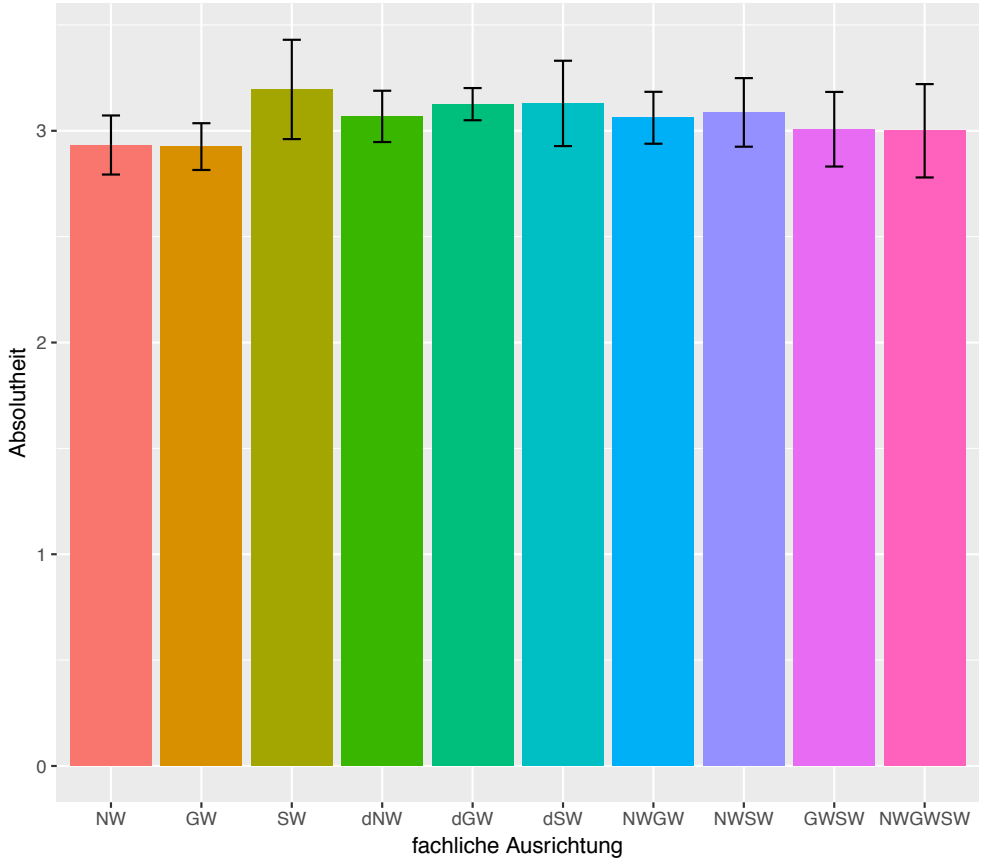


ABBILDUNG I.17: Durchschnittliche Ausprägung in der Überzeugung zur Absolutheit des Wissens nach fachlicher Ausrichtung

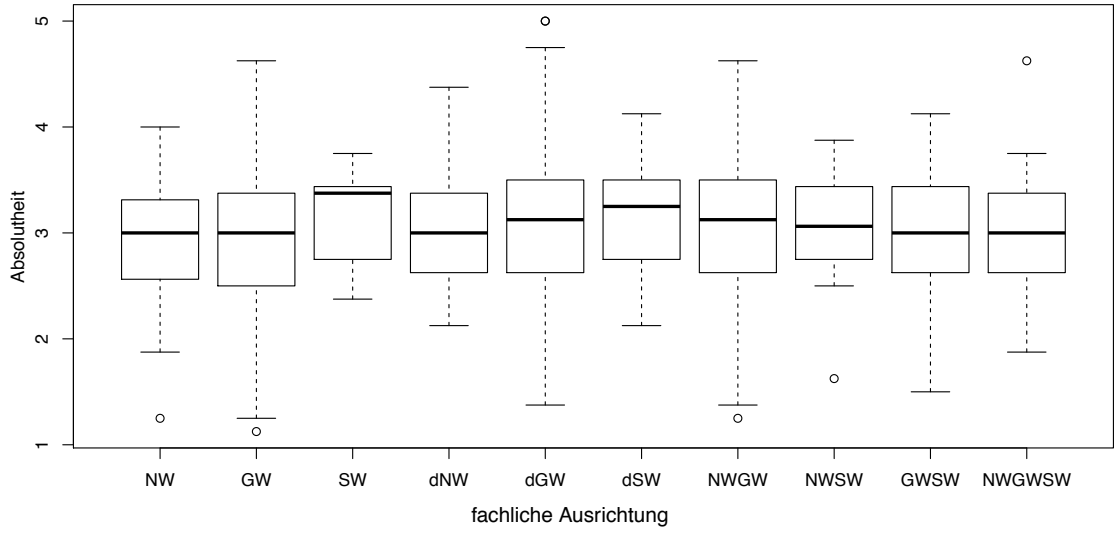


ABBILDUNG I.18: Boxplot der individuellen Skalenmittelwerte zur Absolutheit des Wissens nach fachlicher Ausrichtung

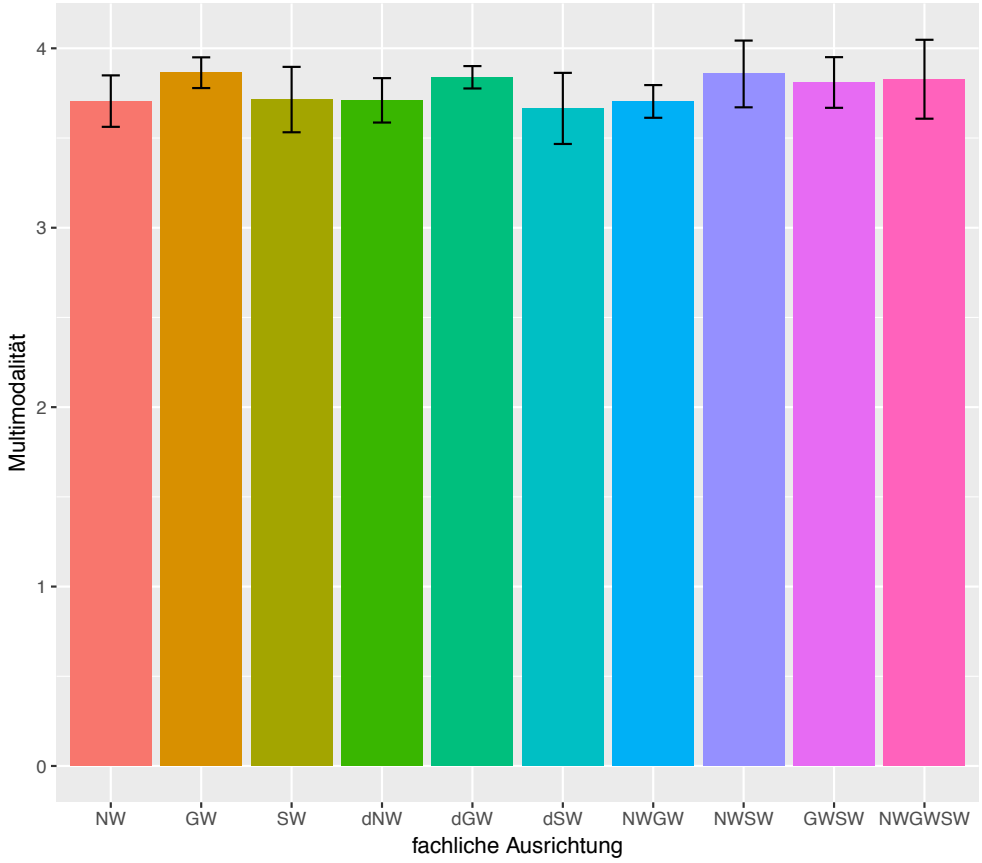


ABBILDUNG I.19: Durchschnittliche Ausprägung in der Überzeugung zur Multimodalität des Wissenserwerbs nach fachlicher Ausrichtung

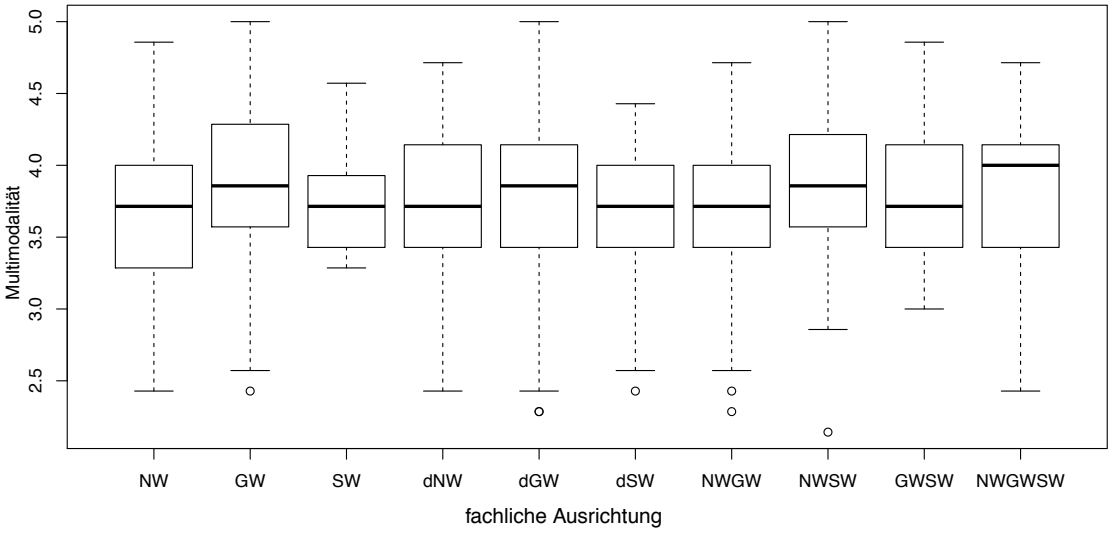


ABBILDUNG I.20: Boxplot der individuellen Skalenmittelwerte zur Multimodalität des Wissenserwerbs nach fachlicher Ausrichtung

# Literaturverzeichnis


Abelson, R. P. (1986). Beliefs are like possessions. *Journal for the Theory of Social Behavior*, *16*, 223–250.

Abrami, P. C., Poulsen, C. & Chambers, B. (2004). Teacher motivation to implement an educational innovation: Factors differentiating users and non-users of cooperative learning. *Educational Psychology*, *24*, 201–216.

Adibelli-Sahin, E. & Bailey, J. M. (2017). Exploring the factors contributing to preservice elementary teachers' epistemological worldviews about teaching science. In G. Schraw, J. Brownlee, L. Olafson & M. VanderVeldt (Hrsg.), *Teachers' personal epistemologies. evolving models for informing practice* (S. 291–322). Charlotte, NC: Information Age Publishing.

AERA, APA & NCME (Hrsg.). (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Alexander, P. A. (1992). Domain knowledge: Evolving themes and emerging concerns. *Educational Psychologist*, *27* (1), 33–51.

Alexander, P. A. & Dochy, F. J. R. C. (1994). Adults' views about knowing and believing. In R. Garner & P. A. Alexander (Hrsg.), *Beliefs about text and instruction with text* (S. 223–244). Hillsdale, NJ: Lawrence Erlbaum Associates.

Andre, T. & Windschitl, M. (2003). Interest, epistemological belief, and intentional conceptual change. In G. M. Sinatra & P. R. Pintrich (Hrsg.), *Intentional conceptual change: Includes bibliographical references and indexes.* Mahwah, N.J.: Erlbaum.

Asendorpf, J. (1999). *Persönlichkeitspsychologie*. Heidelberg: Springer.

Audi, R. (1999). *The cambridge dictionary of philosophy* (2. Aufl.). Cambridge and New York and Melbourne and Cape Town and Singapore and Sao Paulo: Cambridge University Press.

Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (Hrsg.). (2003). *Multivariate analysemethoden: Eine anwendungsorientierte einführung* (10. Aufl.). Berlin: Springer.

Baglin, J. (2014). Improving your exploratory factor analysis for ordinal data: A demonstration using factor. *Practical Assessment, Research & Evaluation*, *19* (5), 1–15. Zugriff auf `http://pareonline.net/getvn.asp?v=19&n=5`

Barbera, J., Perkins, K. K., Adams, W. K. & Wieman, C. E. (2008). Modifying and validating the colorado learning attitudes about science survey for use in chemistry. *Journal of Chemical Education*, *85*, 1435–1439. Zugriff auf `https://www.colorado.edu/sei/class/CLASS%20I.pdf`

Barnard, L., Lan, W. L., Crooks, M. & Paton, V. O. (2008). The relationship between epistemological beliefs and self-regulated learning skills in the online course environment.

*MERLOT Journal of Online Learning and Teaching*, *4* (3), 261–266.

Barth, H., Bhandari, K., Garcia, J., MacDonald, K. & Chase, E. (2014). Preschoolers trust novel members of accurate speakers' groups and judge them favorably. *Quartlery Journal of Experimental Psychology*, *67*, 872–883.

Bartholome, T., Stahl, E., Pieschl, S. & Bromme, R. (2006). What matters in help-seeking? a study of help effectiveness and learner-related factors. *Computers in Human Behavior*, *22*, 113–129.

Barzilai, S. & Zohar, A. (2012). Epistemic thinking in action: Evaluating and integrating online sources. *Cognition and Instruction*, *30* (1), 39–85.

Barzilai, S. & Zohar, A. (2014). Reconsidering personal epistemology as metacognition: A multifaceted approach to the analysis of epistemic thinking. *Educational Psychologist*, *49* (1), 13–35.

Barzilai, S. & Zohar, A. (2016). Epistemic (meta)cognition: Ways of thinking about knowledge and knowing. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 409–424). New York and London: Routledge.

Basto, M. & Pereira, J. M. (2012). An spss r-menu for ordinal factor analysis. *Journal of Statistical Software*, *46* (4), 1–29.

Baumert, J., Bos, W. & Lehmann, R. (Hrsg.). (2000). *Timms/iii dritte internationale mathematik- und naturwissenschaftsstudie - mathematische und physikalische kompetenzen am ende der gymnasialen oberstufe: Mathematische und naturwissenschaftliche grundbildung am ende der pflichtschulzeit* (Bd. 1).

Baxter Magolda, M. B. (1987). The affective dimension of learning: Faculty-student relationships that enhance intellectual development. *College Student Journal*, *21*, 46–58.

Baxter Magolda, M. B. (1992). *Knowing and reasoning in college: Gender-related patterns in students' intellectual development*. San Francisco: Jossey-Bass.

Baxter Magolda, M. B. (2004). Evolution of a constructivist conceptualization of epistemological reflection. *Educational Psychologist*, *39* (1), 31–42.

Baxter Magolda, M. B. (2002). Epistemological reflection: The evolution of epistemological assumptions from age 18 to 30. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 89–102). Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Baxter Magolda, M. B., Creamer, E. G. & Meszaros, P. S. (2010). *Development and assessment of self-authorship: Exploring the concept across cultures*. Sterling, VA: Stylus.

Baxter Magolda, M. B. & Porterfield, W. D. (1985). *Assessing intellectual development: The link between theory and practice*. Alexandria, VA: American College Personnel Association.

Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J. & Esquivel, S. L. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research & Evaluation*, *18* (6), 1–13. Zugriff auf `http://pareonline.net/pdf/v18n6.pdf`

Belenky, M. F., Clinchy, B. M., Goldberger, N. R. & Tarule, J. M. (1986). *Women's ways of knowing: The development of self, voice and mind*. New York: Basic Books.

Bell, P. & Linn, M. C. (2002). Beliefs about science: How does science instruction contribute? In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 321–346). Mahwah,

New Jersey, London: Lawrence Erlbaum Associates.

Bendixen, L. D. (2002). A process model of epistemic belief change. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 191–208). Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Bendixen, L. D. (2016). Teaching for epistemic change in elementary classrooms. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 281–299). New York and London: Routledge.

Bendixen, L. D. & Corkill, A. J. (2011). Personal epistemology change due to experience? a cross-sectional analysis of preservice and practicing teachers. In J. Brownlee, G. Schraw & D. Berthelsen (Hrsg.), *Personal epistemology and teacher education* (S. 100–113). New York, London: Routledge.

Bendixen, L. D., Dunkle, M. E. & Schraw, G. (1994). Epistemological beliefs and reflective judgement. *Psychological Review*, *75* (3), 1595–1600.

Bendixen, L. D. & Feucht, F. C. (Hrsg.). (2010). *Personal epistemology in the classroom: Theory, research, and implications for practice*. Cambridge [u.a.]: Cambridge University Press.

Bendixen, L. D. & Rule, D. C. (2004). An integrative approach to personal epistemology: A guiding model. *Educational Psychologist*, *39* (1), 69–80.

Bendixen, L. D., Schraw, G. & Dunkle, M. E. (1998). Epistemic beliefs and moral reasoning. *The Journal of Psychology*, *132*, 187–200.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107* (2), 238–246.

Berding, F. & Lamping, C. (2014). *Epistemologische überzeugungen als bestandteil der professionellen kompetenz von nlehrkräften und ihre bedeutung für die auswahl und bewertung von lernaufgaben aus schulbüchern des wirtschaftslehreunterrichts: Eine explorative studie* (Bd. 12). München, Mering: Rainer Hampp.

Bernholt, A., Gruber, H. & Moschner, B. (2017). *Wissen und lernen. wie epistemische überzeugungen schule, universität und arbeitswelt beeinflussen*. Münster and New York: Waxmann.

Biggs, J. B. (1987). *The study process questionnaire (spq): Manual*. Hawthorn, Victoria: Australian Council for Educational Research.

Biggs, J. B. (1993). What do inventories of students' learning processes really measure? a theoretical review and clarification. *British Journal of Educational Psychology*, *63*, 3–19.

Biggs, J. B., Kember, D. & Leung, D. Y. P. (2001). The revised two-factor study process questionnaire: R-spq-2f. *British Journal of Educational Psychology*, *71*, 133–149.

Biglan, A. (1973a). The characteristics of subject matter in different academic areas. *Journal of Applied Psychology*, *57* (3), 195–203.

Biglan, A. (1973b). Relationships between subject matter characteristics and the structure and output of university departments. *Journal of Applied Psychology*, *57* (3), 204–213.

Blignaut, S., Mokwinski, B., Paechter, M., Rebmann, K. & Sulimma, M. (2010). Überzeugungen zu wissen und lernen bei deutschen und südafrikanischen lehramtsstudierenden. *Berufsbildung*, *123*, 45–48.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H. & Krathwohl, D. R. (Hrsg.). (1956). *Taxonomy of educational objectives, handbook i: The cognitive domain*. New York: David

McKay Co Inc.

Boekaerts, M., Pintrich, P. R. & Zeidner, M. (2000). *Handbook of self-regulation*. San Diego: Academic Press.

Bollen, K. A. (1989). *Structural equations with latent variables*. Oxford: John Wiley and Sons.

Bollen, K. A. & Long, J. S. (Hrsg.). (1993). *Testing structural equation models* (Bd. 154). Newbury Park: Sage Publications.

Bondy, E., Ross, D., Adams, A., Nowak, R., Brownell, M. & Hoppey, e. a., D. (2007). Personal epistemologies and learning to teach. *Teacher Education and Special Education: The Journal of Teacher Education Division of the Council for Exceptional Children*, *30*, 67–82.

Bortz, J. & Döring, N. (2006). *Forschungsmethoden und evaluation für human- und sozialwissenschaftler: Mit 87 tabellen* (4., überarb. Aufl. Aufl.). Heidelberg: Springer.

Bortz, J. & Schuster, C. (2010). *Statistik für human- und sozialwissenschaftler* (7., vollst. überarb. und erw. Aufl. Aufl.). Berlin: Springer.

Boyes, M. C. & Chandler, M. (1992). Cognitive development, epistemic doubt, and identity formation in adolescence. *Journal of Youth and Adolescence*, *21* (3), 277–304.

Bråten, I. (2016). Epistemic cognition interventions. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 360–371). New York and London: Routledge.

Bråten, I. & Olaussen, B. S. (2005). Profiling individual differences in student motivation: A longitudinal cluster-analytic study in different academic contexts. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *30* (3), 359–396.

Bråten, I. & Strømsø, H. I. (2004). Epistemological beliefs and implicit theories of intelligence as predictors of achievement goals. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *29* (4), 371–388.

Bråten, I. & Strømsø, H. I. (2006). Effects of personal epistemology on the understanding of multiple texts. *Reading Psychology*, *27*, 457–484.

Bråten, I. & Strømsø, H. I. (2010). Effects of task instruction and personal epistemology on the understanding of multiple texts about climate change. *Discourse Processes*, *47* (1), 1–31.

Bråten, I., Strømsø, H. I. & Samuelstuen, M. S. (2008). Are sophisticated students always better? the role of topic-specific personal epistemology in the understanding of multiple expository texts. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *33* (4), 814–840.

Brickhouse, N. W. (1989). The teaching of the philosophy of science in secondary classrooms. *International Journal of Science Education*, *11*, 437–449.

Briggs, S. R. & Cheek, J. M. (1986). The role pf factor analysis in the development and evaluation of personality scales. *Journal of Personality*, *54* (1), 106–148.

Bromme, R. (2010). Epistemological beliefs are standards for adaptive learning: A functional theory about epistemological beliefs and metacognition. *Metacognition and Learning*, *5* (1), 7–26.

Bromme, R., Kienhues, D. & Stahl, E. (2008). Knowledge and epistemological beliefs: An intimate but complicate relationship. In M. S. Khine (Hrsg.), *Knowing, knowledge and*

*beliefs* (S. 423–444). Dordrecht: Springer.

Bromme, R., Pieschl, S. & Stahl, E. (2014). Epistemological beliefs and students' adaptive perception of task complexity. In S. Krolak-Schwerdt, S. Glock & M. Böhmer (Hrsg.), *The future of educational research* (Bd. 3, S. 123–151). Rotterdam: Sense Publishers.

Bromme, R. & Stahl, E. (2005). Is a hypertext a book or a space? the impact of different introductory metaphors on hypertext construction. *Computers & Education*, *44* (2), 115–133.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (Second edition Aufl.). New York: Guilford Publications.

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, *36*, 111–150.

Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Hrsg.), *Testing structural equation models* (Bd. 154, S. 136–162). Newbury Park: Sage Publications.

Brownlee, J., Ferguson, L. E. & Ryan, M. (2017). Changing teachers' epistemic cognition: A new conceptual framework for epistemic reflexivity. *Educational Psychologist*, *52* (4), 242–252.

Brownlee, J., Johansson, E., Walker, S. & Scholes, L. (2017). *Teaching for active citizenship: Personal epistemology and practices in early education classrooms*. New York, NY: Routledge.

Brownlee, J., Schraw, G. & Berthelsen, D. (Hrsg.). (2011). *Personal epistemology and teacher education*. New York, London: Routledge.

Brownlee, J., Schraw, G. & Berthelsen, D. (2013). *Personal epistemology and teacher education* (1. Aufl., Bd. 61). New York [u.a.]: Routledge.

Brownlee, J., Schraw, G., Walker, S. & Ryan, M. (2016). Changes in preservice teachers' personal epistemologies. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 300–317). New York and London: Routledge.

Bryant, F. B. & Yarnold, P. R. (2001). Principal component analysis and exploratory and confirmatory factor analysis. In L. G. Grimm & P. R. Yarnold (Hrsg.), *Reading and understanding more multivariate statistics* (S. 99–136). Washington: American Psychological Association.

Buehl, M. M. (2008). Assessing the multidimensionality of students' epistemic beliefs across diverse cultures. In M. S. Khine (Hrsg.), *Knowing, knowledge and beliefs* (S. 65–112). Dordrecht: Springer.

Buehl, M. M. & Alexander, P. A. (2001). Beliefs about academic knowledge. *Educational Psychology Review*, *13* (4), 385–418.

Buehl, M. M. & Alexander, P. A. (Hrsg.). (2004). *Modeling the relations between domain-specific epistemological beliefs, motivation, and task performance, paper presented at the annual meeting of the american educational research association.* San Diego, CA.

Buehl, M. M. & Alexander, P. A. (2005). Motivation and performance differences in students' domain-specific epistemological belief profiles. *American Educational Research Journal*, *42* (4), 697–726.

Buehl, M. M. & Alexander, P. A. (2006). Examining the dual nature of epistemological

beliefs. *International Journal of Educational Research*, *45* (1/2), 28–42.

Buehl, M. M., Alexander, P. A. & Murphy, P. K. (2002). Beliefs about schooled knowledge: Domain specific or domain general? *Epistemological Development and Its Impact on Cognition in Academic Domains*, *27* (3), 415–449.

Buehl, M. M. & Fives, H. (2009). Exploring teachers' beliefs about teaching knowledge: Where does it come from? does it change? *The Journal of Experimental Education*, *77*, 367–407.

Buehl, M. M. & Fives, H. (2016). The role of epistemic cognition in teacher learning praxis. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 247–264). New York and London: Routledge.

Bühner, M. (2011). *Einführung in die test- und fragebogenkonstruktion* (3., aktualisierte und erw. Aufl. Aufl.). München: Pearson.

Burr, J. E. & Hofer, B. K. (2002). Personal epistemology and theory of mind: Deciphering young children's beliefs about knowledge and knowing. *Folk Epistemology*, *20* (2–3), 199–224.

Byrne, B. (2010). *Structural equation modeling using amos. basic concepts, application, and programming* (Bd. 2). New York: Routledge.

Cano, F. (2005). Epistemological beliefs and approaches to learning: Their change through secondary school and their influence on academic performance. *British Journal of Educational Psychology*, *75* (2), 203–221.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*, 245–276.

Chai, C. S., Khine, M. S. & Teo, T. (2006). Epistemological beliefs on teaching and learning: A survey among pre-service teachers in singapore. *Educational Media International*, *43*, 285–298.

Chan, K. & Elliott, R. G. (2000). Exploratory study of epistemological beliefs of hong kong teacher education students: Resolving conceptual and empirical issues. *Asia-Pacific Journal of Teacher Education*, *28* (3), 225–234.

Chan, K. & Elliott, R. G. (2002). Exploratory study of hong kong teacher education students' epistemological beliefs: Cultural perspectives and implications on beliefs research. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *27* (3), 392–414.

Chandler, M. J. (1975). Relativism and the problem of epistemological loneliness. *Human Development*, *18*, 171–180.

Chandler, M. J. (1987). The othello effect: Essay on the emergence and eclipse of skeptical doubt. *Human Development*, *30*, 137–159.

Chandler, M. J., Boyes, M. & Ball, L. (1990). Relativism and stations of epistemic doubt. *Journal of Experimental Child Psychology*, *50*, 370–395.

Chandler, M. J., Hallett, D. & Sokol, B. W. (2002). Competing claims about competing knowledge claims. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 145–168). Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Charalambous, C. Y., Panaoura, A. & Phillippou, G. (2009). Using the history of mathematics

to induce changes in preservice teachers' beliefs and attitudes: Insights from evaluating a teacher education program. *Educational Studies of Mathematics*, *71*, 161–180.

Chinn, C. A., Bruckland, L. A. & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. *Educational Psychologist*, *46* (3), 141–167.

Chinn, C. A. & Rinehart, R. W. (2016). Epistemic cognition and philosophy: Developing a new framework for epistemic cognition. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 460–478). New York and London: Routledge.

Chinn, C. A., Rinehart, R. W. & Buckland, L. A. (2014). Epistemic cognition and evaluating information: Applying the air model of epistemic cognition. In R. N. Rapp & B. J. L. G (Hrsg.), *Processing inaccurate information* (S. 425–453). Cambridge: MIT Press.

Christmann, U. & Groeben, N. (1999). Psychologie des lesens. In B. Franzmann, K. Hasemann & D. Löffler (Hrsg.), *Handbuch lesen* (S. 145–223). München: Saur.

Clarebout, G., Elen, J., Luyten, L. & Bamps, H. (2001). Assessing epistemological beliefs: Schommer's questionnnaire revisited. *Educational Research and Evaluation*, *7*, 53–77.

Cole, R. P. & Goetz, E. T. W. V. (2000). Epistemological beliefs of underprepared college students. *Journal of College Reading and Learning*, *31* (1), 60–72.

Conley, A. M., Pintrich, P. R., Vekiri, I. & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *29* (2), 186–204.

Corno, L. (1993). The best-laid plans: Modern conceptions of volition and educational research. *Educational Researcher*, *22* (2), 14–22.

Costello, A. B. & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, *10* (7), 1–9. Zugriff auf `http://pareonline.net/pdf/v10n7.pdf`

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334.

Cronbach, L. J. & Mehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302.

Dahl, T. I., Bals, M. & Turi, A. L. (2005). Are students' beliefs about knowledge and learning associated with their reported use of learning strategies? *British Journal of Educational Psychology*, *75* (2), 257–273.

Daniels, D. H. & Shumow, L. (2003). Child development and classroom teaching: A review og the literature and implications for educational teachers. *Applied Developmental Psychology*, *23*, 495–526.

Dann, H.-D. (1994). Pädagogisches verstehen. In K. Reusser & M. Reusser-Weyeneth (Hrsg.), *Psychologischer prozeß und didaktische aufgabe* (S. 163–182). Bern: Huber.

DeBacker, T. K. & Crowson, H. M. (2006). Influences on cognitive engagement: Epistemological beliefs and need for closure. *The British journal of educational psychology*, *76* (Pt 3), 535–551.

DeBacker, T. K. & Crowson, H. M. (2008). Measuring need for closure in classroom learners.

*Epistemological Development and Its Impact on Cognition in Academic Domains*, *33* (4), 711–732.

de Corte, E., Op't Eynde, P. & Verschaffel, L. (2002). Knowing what to belief: The relevanceof students' mathematical beliefs for mathematics education. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 297–320). Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Deniz, H. (2011). Examination of changes in prospective elementary teachers' epistemological beliefs in science and exploration of factors mediating that change. *Journal of Science Education and Technology*, *20*, 750–760.

Depaepe, F., de Corte, E. & Verschaffel, L. (2016). Mathematical epistemological beliefs. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 147–164). New York and London: Routledge.

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. Lexington, MA: Heath.

Dewey, J. (1938). *Logic: The theory of inquiry*. Troy, MA: Holt, Rinehart und Winston.

DiStefano, C. & Motl, R. W. (2009). Personality correlates of method effects due to negatively worded items on the rosenberg self-esteem scale. *Personality and Individual Differences*, *46*, 309–313.

Dole, J. A. & Sinatra, G. M. (1998). Reconceptalizing change in the cognitive construction of knowledge. *Educational Psychologist*, *33* (2-3), 109–128.

Donald, J. G. (1990). University professors' views of knowledge and validation processes. *Journal of Educational Psychology*, *82* (2), 242–249.

Donald, J. G. (1995). Disciplinary differences in knowledge validation. In N. Hativa & M. Marincovich (Hrsg.), *Disciplinary differences in teaching and learning: Implications for practice* (S. 7–18). San Francisco: Jossey-Bass.

Döring, N. & Bortz, J. (2016). *Forschungsmethoden und evaluation in den sozial- und humanwissenschaften* (5. vollst. überarb., aktualisierte und erw. Aufl. Aufl.). Berlin: Springer.

Duell, O. K. & Schommer-Aikins, M. (2001). Measures of people's beliefs about knowledge and learning. *Educational Psychology Review*, *13*, 419–449.

Dweck, C. & Legett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95* (2), 256–273.

Elby, A. (2001). On the substance of a sophisticated epistemology. *Science Education*, *85* (5), 554–567.

Elby, A. & Hammer, D. (2010). Epistemological resources and framing: A cognitive framework for helping teachers interpret and respond to their students' epistemologies. In L. D. Bendixen & F. C. Feucht (Hrsg.), *Personal epistemology in the classroom* (S. 409–434). Cambridge [u.a.]: Cambridge University Press.

Elby, A., Macrander, C. & Hammer, D. (2016). Epistemic cognition in science. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 113–127). New York and London: Routledge.

Elder, A. D. (2002). Characterizing fifth grade students' epistemological beliefs in science. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 347–363). Mahwah, New

Jersey, London: Lawrence Erlbaum Associates.

Elen, J., Stahl, E., Bromme, R. & Clarebout, G. (2011). *Links between beliefs and cognitive flexibility: Lessons learned*. Dordrecht: Springer.

Enders, C. K. (2010). *Applied missing data analysis*. New York, NY: Guilford Press.

Estes, D., Chandler, M., Horvath, K. J. & Backus, D. W. (2003). American and british college students' epistemological beliefs about research on psychological and biological development. *Applied Developmental Psychology*, *23*, 625–642.

Fabrigar, L. R. & Wegener, D. T. (2012). *Exploratory factor analysis*. Oxford and New York: Oxford University Press.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4* (3), 272–299.

Ferguson, L. E. & Bråten, I. (2013). Student profiles of knowledge and epistemic beliefs: Changes and relations to multiple-text comprehension. *Learning and Instruction*, *25*, 49–61.

Ferguson, L. E., Bråten, I. & Strømsø, H. I. (2012). Epistemic cognition when students read multiple documents containing conflicting scientific evidence: A think-aloud study. *Learning and Instruction*, *22* (2), 103–120.

Field, A., Miles, J. & Field, Z. (2012). *Discovering statistics using r* (1. publ Aufl.). Los Angeles, Calif.: Sage.

Fisseni, H. J. (Hrsg.). (1990). *Lehrbuch der psychologischen diagnostik*. Göttingen: Hogrefe.

Fisseni, H. J. (Hrsg.). (1997). *Lehrbuch der pädagogischen diagnostik*. Göttingen: Hogrefe.

Fives, H. (2011). One preservice teacher's developing epistemology about teaching and the explicit connection of those beliefs to future practice. In J. Brownlee, G. Schraw & D. Berthelsen (Hrsg.), *Personal epistemology and teacher education* (S. 114–128). New York, London: Routledge.

Fives, H. & Buehl, M. M. (2010). Teachers' articulation of beliefs about teaching knowledge: Conceptualizing a belief framework. In L. D. Bendixen & F. C. Feucht (Hrsg.), *Personal epistemology in the classroom* (S. 470–515). Cambridge [u.a.]: Cambridge University Press.

Fives, H. & Buehl, M. M. (2012). Spring cleaning for the messy construct of teachers' beliefs: What are they? which have been examined? what can they tell us? In K. R. Harris, S. Graham & T. Urdan (Hrsg.), *Apa educational psychology handbook, vol. 2, individual differences and cultural and contextual factors* (Bd. 2, S. 471–499). American Psychological Association.

Fives, H. & Buehl, M. M. (2017). The functions of beliefs: Teachers' personal epistemology on the pinning blocks. In G. Schraw, J. Brownlee, L. Olafson & M. VanderVeldt (Hrsg.), *Teachers' personal epistemologies. evolving models for informing practice* (S. 25–54). Charlotte, NC: Information Age Publishing.

Fives, H. & Gill, M. G. (2014). *International handbook of research on teachers' beliefs*. New York and London: Routledge.

Flavell, J. H. (1971). Metacognitive aspects of problem solving. In L. B. Resnick (Hrsg.), *The*

*nature of intelligence* (S. 82–121). Hillsdale: Erlbaum.

Flavell, J. H. (1979). Metacognition and cognitive monitoring. *American Psychologist*, *34*, 906–911.

Flavell, J. H. & Wellman, H. M. (1977). Metamemory. In R. V. Kail & J. W. Hagen (Hrsg.), *Perspectives on the development of memory and cognition.* Hillsdale: Erlbaum.

Franco, G. M., Muis, K. R., Kendeou, P., Ranellucci, J., Sampasivam, L. & Wang, X. (2012). Examining the influences of epistemic beliefs and knowledge representations on cognitive processing and conceptual change when learning physics. *Learning and Instruction*, *22* (1), 62–77.

Frederiksen, N. (1984). Implications of cognitive theory for instruction in problem solving. *Review of Educational Research*, *54* (3), 363–407.

Friedrich, H. F. & Mandl, H. (1992). Lern- und denkstrategien: Ein problemaufriss. In H. Mandl & H. F. Friedrich (Hrsg.), *Lern- und denkstrategien* (S. 3–54). Göttingen: Hogrefe.

Gallagher, L. P. (2002). *Adolescent ways of knowing: Profiles of interdependence in epistemic beliefs, contexts, and practices* (Dissertation). Stanford University, US.

Galloti, K., McVicker Clinchy, B., Ainsworth, K., Lavin, B. & Manfield, A. (1999). A new way of assessing ways of knowing: The attitudes towards thinking and learning survey (attls). *Sex Roles*, *40*, 745–766.

Garner, R. & Alexander, P. A. (Hrsg.). (1994). *Beliefs about text and instruction with text*. Hilldale: Erlbaum.

Garrett-Ingram, C. (Hrsg.). (1997). *Something to believe in: The relationship between epis- temological beliefs and study strategies, paper presented at the american educational research association.* Chicago, IL.

Gerber, J. (2004). *Intergenerationale transmission epistemologischer überzeugungen. entwicklung eines erhebungsinstruments und befunde zur genese wissensbezogener vorstellungen* (Dissertation). Universität Bielefeld.

Gholami, K. (2017). Understanding the epistemic nature of teachers' reasoning behind their practices from an aristotelian perspective. In G. Schraw, J. Brownlee, L. Olafson & M. VanderVeldt (Hrsg.), *Teachers' personal epistemologies. evolving models for informing practice* (S. 141–168). Charlotte, NC: Information Age Publishing.

Gill, M. G., Ashton, P. T. & Algina, J. (2004). Changing preservice teachers' epistemological beliefs about teaching and learning in mathematics: An intervention study. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *29* (2), 164–185.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.

Goldman, A. I. & Blanchard, T. (2012). Social epistemology. *Oxford Bibliographies Online*. Zugriff auf `www.oxfordbibliographies.com`

Gopnik, A. & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind and Language*, *7*, 145–171.

Gottlieb, E. (2007). Learning how to believe: Epistemic development in cultural context. *Journal of the Learning Sciences*, *16* (1), 5–35.

Greene, J. A., Azevedo, R. & Torney-Purta, J. (2008). Modeling epistemic and ontological cognition: Philosophical perspectives and methodological directions. *Educational Psychologist*, *43*, 142–160.

Greene, J. A., Muis, K. R. & Pieschl, S. (2010). The role of epistemic beliefs in students' self-regulated learning with computer-based learning environments: Conceptual and methodological issues. *Educational Psychologist*, *45* (4), 245–257.

Greene, J. A., Sandoval, W. A. & Bråten, I. (Hrsg.). (2016). *Handbook of epistemic cognition*. New York and London: Routledge.

Grossnickle, E. M., Alexander, P. A. & List, A. (2017). The argument for epistemic competence. In A. Bernholt, H. Gruber & B. Moschner (Hrsg.), *Wissen und lernen. wie epistemische überzeugungen schule, universität und arbeitswelt beeinflussen* (S. 255–270). Münster and New York: Waxmann.

Gruber, H. & Moschner, B. (2003). *Begabung, lernen, wissen: Epistemologische überzeugungen.*

Guthrie, V. L., King, P. M. & Palmer, C. J. (2000). Higher education and reducing prejudice: Research on cognitive capabilities underlying tolerance. *Diversity Digest*, *4* (3), 10–23.

Guttman, L. (1954). Some necessary conditions for common factor analysis. *Psychometrika*, *19* (2), 149–161.

Güven, G., Sülün, Y. & Cam, A. (2011). The examination of elementary preservice teachers' reflective diaries and epistemological beliefs in science laboratory. *Teaching in Higher Education*, *19*, 895–907.

Guzzetti, B. J., Snyder, T. E., Glass, G. V. & Gamas, W. S. (1993). Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. *Reading Research Quarterly*, *28*, 117–155.

Häder, M. (2014). *Delphi-befragungen: Ein arbeitsbuch* (3. Aufl. Aufl.). Wiesbaden: Springer VS.

Haerle, F. C. (Hrsg.). (2006). *Personal epistemologies of fourth graders: Their belief about knowledge and knowing*. Didaktisches Zentrum: Oldenburg.

Haerle, F. C. & Bendixen, L. D. (2008). Personal epistemology in elementary classrooms: A conceptual comparison of germany and the united states and a guide for future cross-cultural research. In M. S. Khine (Hrsg.), *Knowing, knowledge and beliefs* (S. 151–176). Dordrecht: Springer.

Hallett, D., Chandler, M. J. & Krettenauer, T. (2002). Disentangling the course of epistemic development: Parsing knowledge by epistemic content. *New Ideas in Psychology*, *20*, 285–307.

Halloun, I. & Hestenes, D. (1998). Interpreting vass dimensions and profiles. *Science & Education*, *7* (6), 553–577.

Hammer, D. (1994). Epistemological beliefs in introductory physics. *Cognition and Instructions*, *12*, 151–183.

Hammer, D. & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 169–190). Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Hammer, D. & Elby, A. (2003). Tapping epistemological resources for learning physics. *The*

*Journal of the Learning Sciences*, *12* (1), 53–90.

Hammrich, P. L. (1998). Cooperative controversy challenges elementary teacher candidates' conceptions of the nature of science. *Journal of Elementary Science Education*, *10*, 50–65.

Hartig, J., Frey, A. & Jude, N. (2012). Validität. In H. Moosbrugger (Hrsg.), *Testtheorie und fragebogenkonstruktion* (S. 143–171). Berlin: Springer.

Hartig, J., Frey, A. & Jude, N. (im Druck). Validität von testwertinterpretationen. In H. Moosbrugger (Hrsg.), *Testtheorie und fragebogenkonstruktion.* Berlin: Springer.

Hartinger, A., Kleickmann, T. & Hawelka, B. (2006). Der einfluss von lehrervorstellungen zum lernen und lehren auf die gestaltung des unterrichts und auf motivationale schülervariablen. *Zeitschrift für Erziehungswissenschaft*, *9* (1), 110–126.

Hashweh, M. Z. (1996). Effects of science teachers' epistemological beliefs in teaching. *Journal of Research in Science Teaching*, *33* (1), 47–63.

Helmke, A. (2005). *Unterrichtsqualität: Erfassen, bewerten, verbessern* (4. Aufl.). Seelze: Kallmeyer-Verlag.

Hofer, B. K. (1999). Instructional context in the college mathematics classroom: Epistemological beliefs and student motivation. *Journal of Staff, Program and Organization Development*, *16* (2), 73–82.

Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *25* (4), 378–405.

Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Educational Psychology Review*, *13* (4), 353–384.

Hofer, B. K. (2002). Personal epistemology as a psychological and educational construct: An introduction. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 3–14). Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Hofer, B. K. (2004a). Epistemological understanding as a metacognitive process: Thinking aloud during online searching. *Educational Psychologist*, *39* (1), 43–55.

Hofer, B. K. (2004b). Exploring the dimensions of personal epistemology in differing classroom contexts: Student interpretations during the first year of college. *Contemporary Educational Psychology*, *29* (2), 129–163.

Hofer, B. K. (2005). The legacy and the challenges: Paul pintrich's contributions to personal epistemology. *Educational Psychologist*, *40* (2), 95–105.

Hofer, B. K. (2006a). Domain specificity of personal epistemology: Resolved questions, persistent issues, new models. *International Journal of Educational Research*, *45* (1–2), 85–95.

Hofer, B. K. (2006b). Beliefs about knowledge and knowing: Intregrating domain specificity and domain generality: A response to muis, bendixen, and haerle (2006). *Educational Psychology Review*, *18* (1), 67–76.

Hofer, B. K. (Hrsg.). (2007). *Learning strategies and epistemic beliefs: Cultural influences in japan and the us. paper presented at the american educational research association.* Chicago.

Hofer, B. K. (2008). Personal epistemology and culture. In M. S. Khine (Hrsg.), *Knowing, knowledge and beliefs* (S. 3–22). Dordrecht: Springer.

Hofer, B. K. (2016). Epistemic cognition as a psychological construct: Advancements and challenges. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 19–38). New York and London: Routledge.

Hofer, B. K. & Bendixen, L. D. (2012). Personal epistemology: Theory, research, and future directions. In K. R. Harris, S. Graham & T. Urdan (Hrsg.), *Apa educational psychology handbook, vol. 3, application to learning and teaching* (S. 227–256). Washington, D.C.: American Psychological Association.

Hofer, B. K. & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, *67* (1), 88–140.

Hofer, B. K. & Pintrich, P. R. (Hrsg.). (2002). *Personal epistemology: The psychology of beliefs about knowledge and knowing*. Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Hofer, B. K. & Sinatra, G. M. (2010). Epistemology, metacognition, and self-regulation: Musings on an emerging field. *Metacognition and Learning*, *5* (1), 113–120. doi: 10.1007/s11409-009-9051-7

Homburg, C. & Baumgartner, H. (1995). Beurteilung von kausalmodellen: Bestandsaufnahme und anwendungsempfehlungen. *Marketing - Zeitschrift für Forschung und Praxis*, *17* (3), 162–176.

Homburg, C. & Giering, A. (1996). Konzeptualisierung und operationalisierung komplexer konstrukte: Ein leitfaden für die marketingforschung. *Marketing - Zeitschrift für Forschung und Praxis*, *18* (1), 5–24.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30* (2), 179–185.

Hu, L. & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3* (4), 424–453.

Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6* (1), 1–55.

Hubbard, R. & Allen, S. J. (1987). An empirical comparison of alternative methods for principal component extraction. *Journal of Business Research*, *15* (2), 173–190.

Jacobson, M. J. & Jehng, J. C. (1999). *Epistemological belief instrument: Scales and items.* Zugriff auf `https://www.rug.nl/research/portal/files/2847842/appendix.pdf`

Jacobson, M. J. & Spiro, R. J. (1995). Hypertext learning environments, cognitive flexibility, and the transfer of complex knowledge: An empirical investigation. *Journal of Educational Computing Research*, *12*, 301–333.

Jehng, J.-C. J., Johnson, S. D. & Anderson, R. C. (1993). Schooling and students' epistemological beliefs about learning. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *18* (1), 23–35.

Jonassen, D. H., Marra, R. M. & Palmer, B. (2004). Epistemological development: An implicit entailment of constructivist learning environments. In N. M. Seel & S. Dikjstra (Hrsg.),

*Curriculums, plans and processes in instructional design: International perspectives* (S. 75–88). Mahwah, N.J.: Lawrence Erlbaum.

Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Hrsg.), *Testing structural equation models* (S. 294–317). Newbury Park: Sage.

Kaiser, H. F. & Dickman, K. (1959). Analytic determination of common factors. *American Psychologist*, *14*, 425.

Kaiser, H. F. & Rice, J. (1974). Little jiffy, mark iv. *Educational and Psychological Measurement*, *34* (1), 11–117.

Kane, M. T. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, *38*, 319–342.

Kang, N. (2008). Learning to teach science: Personal epistemologies, teaching goals, and practices of teaching. *Teaching and Teacher Education*, *24*, 478–498.

Karabenick, S. A. (2005). Culture and personal epistemology: U.s. and middle eastern students' beliefs about scientific knowledge and knowing. *Social Psychology of Education*, *8* (4), 375–394.

Kardash, C. M. & Noel, L. K. (2000). How organizational signals, need for cognition, and verbal ability affect text recall and recognition. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *25* (3), 317–331.

Kardash, C. M. & Scholes, R. J. (1996). Effects of preexisiting beliefs, epistemological beliefs, and need for cognition on interpretation of controversial issues. *Journal of Educational Psychology*, *88* (2), 260–271.

Kelly, G. J. (2016). Methodological considerations for the study of epistemic cognition in practice. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 393–408). New York and London: Routledge.

Kember, D., Leung, D., Jones, A., Loke, A. J., MacKay, J., Sinclair, K., … Young, E. (2000). Development of a questionnaire to measure the level of reflective thinking. *Assessment and Evaluation in Higher Education*, *25*, 381–389.

Khine, M. S. (Hrsg.). (2008). *Knowing, knowledge and beliefs: Epistemological studies across diverse cultures*. Dordrecht: Springer.

Kienhues, D., Bromme, R. & Stahl, E. (2008). Changing epistemological beliefs: The unexpected impact of a short-term intervention. *British Journal of Educational Psychology*, *78*, 545–565.

Kienhues, D., Stadtler, M. & Bromme, R. (2011). Dealing with conflicting or consistent medical information on the web: When expert information breeds laypersons' doubts about experts. *Learning and Instruction*, *21* (2), 193–204.

King, P. M. (1977). *The development of reflective judgment and formal operational thinking in adolescents and young adults* (Dissertation Abstracts International, 38, 7233A). University of Minnesota.

King, P. M. & Baxter Magolda, M. B. (2005). A developmental model of intercultural maturity. *Journal of College Student Development*, *46*, 571–592.

King, P. M. & Kitchener, K. S. (Hrsg.). (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults* (1st ed.

Aufl.). San Francisco: Jossey-Bass Publ.

King, P. M. & Kitchener, K. S. (2002). The reflective judgment model: Twenty years of research on epistemic cognition. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 37–61). Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

King, P. M. & Kitchener, K. S. (2004). Reflective judgment: Theory and research on the development of epistemic assumptions through adulthood. *Educational Psychologist*, *39* (1), 5–18.

King, P. M., Kitchener, K. S., Davison, M. L., Parker, C. A. & Wood, P. K. (1983). The justification of beliefs inn young adults: A longitudinal study. *Human Development*, *26*, 106–116.

King, P. M., Wood, P. K. & Mines, R. A. (1990). Critical thinking among college and graduate students. *The Review of Higher Education*, *13*, 167—186.

Kitchener, K. S. (1983). Cognition, metacognition, and epistemic cognition: A three-level model of cognitive processing. *Human Development*, *26*, 106–116.

Kitchener, K. S. & King, P. M. (1981). Reflective judgment: Concepts of justification and their relationship to age and education. *Journal of Applied Developmental Psychology*, *2*, 89–116.

Kitchener, K. S., King, P. M., Wood, P. K. & Davison, M. L. (1989). Sequentiality and consistency in the development of reflective judgment: A six-year longitudinal study. *Journal of Applied Developmental Psychology*, *10*, 73–95.

Kitchener, K. S., Lynch, C. L., Fischer, K. W. & Wood, P. K. (1993). Developmental range of reflective judgment: The effect of contextual support and practice on developmental stage. *Developmental Psychology*, *29* (5), 893–906.

Kitchener, K. S. & Wood, P. K. (1987). Development of concepts of justification in german university students. *International Journal for Behavioral Development*, *10*, 171–185.

Kline, P. (Hrsg.). (1994). *An easy guide to factor analysis*. London: Routledge.

Kline, R. B. (2013). Exploratory and confirmatory factor analysis. In Y. Petscher & C. Schatsschneider (Hrsg.), *Applied quantitative analysis in the social sciences* (S. 171–207). New York: Routledge.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (Fourth edition Aufl.). New York, London: The Guilford Press.

Klopp, E. (2014). *Die struktur epistemologischer überzeugungen: Empirische und theoretische analysen* (Dissertation). Universität des Saarlandes, Saarbrücken.

Köller, O. (2001). Mathematical world views and achievement in advanced mathematics in germany: Findings from timss population 3. *Studies in Educational Evaluation*, *27* (1), 65–78.

Köller, O., Baumert, J. & Neubrand, J. (2000). Epistemologische überzeugungen und fachverständnis im mathematik- und physikunterricht. In J. Baumert, W. Bos & R. Lehmann (Hrsg.), *Timss/iii dritte internationale mathematik- und naturwissenschaftsstudie mathematische und naturwissenschaftliche bildung am ende der schullaufbahn* (S. 229–269). Wiesbaden: VS Verlag für Sozialwissenschaften.

Krettenauer, T. (2005). Die erfassung des entwicklungsniveaus epistemologischer überzeugungen und das problem der übertragbarkeit von interviewverfahren in standardisierte fragebogenmethoden. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *37* (2), 69–79.

Kuhn, D. (1991). *The skills of argument* (1. publ Aufl.). Cambridge: Cambridge University Press.

Kuhn, D. (1999). Metacognitive development. In L. Balter & C. Tamis-Le Monda (Hrsg.), *Child psychology: A handbook of contemporary issues* (S. 258–286). Philadelphia: Psychology Press.

Kuhn, D. (1999a). A developmental model of critical thinking. *Educational Researcher*, *28* (2), 16–26.

Kuhn, D. (1999b). Metacognitive development. In L. Balter & C. S. Tamis-LeMonda (Hrsg.), *Child psychology: A handbook of contemporary issues* (S. 258-286). Philadelphia: Psychology Press.

Kuhn, D. (2000). Metacognitive development. *Current Directions in Psychological Science*, *9*, 178–181.

Kuhn, D., Cheney, R. & Weinstock, M. (2000). The development of epistemological understanding. *Cognitive Development*, *15* (3), 309–328.

Kuhn, D., Iordanou, K., Pease, M. & Wirkala, C. (2008). Beyond control of variables: What needs to develop to achieve skilled scientific thinking? *Cognitive Development*, *23* (4), 435–451.

Kuhn, D. & Park, S.-H. (2005). Epistemological understanding and the development of intellectual values. *International Journal of Educational Research*, *43* (3), 111–124.

Kuhn, D. & Pearsall, S. (2000). Developmental origins of scientific thinking. *Journal of Cognition and Development*, *1* (1), 113–129.

Kuhn, D. & Weinstock, M. (2002). What is epistemological thinking and why does it matter. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 121–144). Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Kuhn, D., Zillmer, N., Crowell, A. & Zavala, J. (2013). Developing norms and argumentation: Metacognitive, epistemological, and social dimensions of developing argumentive competence. *Cognition and Instruction*, *31* (4), 456–496.

Kunter, M. et al. (Hrsg.). (2003). *Pisa 2000: Dokumentation der erhebungsinstrumente*. Berlin: Max-Planck-Institut für Bildungsforschung.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In A. Musgrave (Hrsg.), *Criticism and the growth of knowledge* (S. 91–196). Cambridge: Cambridge University Press.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching. *American Educational Research Journal*, *27*, 29–63.

Lan, W. Y., Bremer, R., Stevens, T. & Mullen, G. (Hrsg.). (2004, April). *Self-regulated learning in the online environment, paper presented at the annual meeting of the american educational research association.* San Diego, CA.

Lance, C. E., Butts, M. M. & Michels, L. C. (2006). The sources of four commonly reported

cutoff criteria: What did they really say? *Organizational Research Methods*, *9* (2), 202–220.

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L. & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching*, *39* (6), 497–521.

Lederman, N. G. & Zeidler, D. (1987). Science teachers' conceptions of the nature of science: Do they really influence teaching behavior. *Science Education*, *7*, 721–734.

Lee, C. D., Goldman, S. R., Levine, S. & Magliano, J. (2016). Epistemic cognition in literary reasoning. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 165–183). New York and London: Routledge.

Levitt, K. E. (2001). An analysis of elementary teachers' beliefs regarding the teaching and learning of science. *Science Education*, *86* (1), 1–22.

Lin, C. H. (2002). Effects of computer graphics types and epistemological beliefs on students' learning of mathematical concepts. *Journal of Educational Computing Research*, *27*, 265–274.

Little, R. J. & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.

Little, R. J. & Rubin, D. B. (2002). *Statistical analysis with missing data* (2. ed. Aufl.). Hoboken, NJ: Wiley-Interscience.

Lonka, K. & Lindblom-Ylänne, S. (1996). Epistemologies, conceptions of learning, and study practices in medicine and psychology. *Higher Education*, *31*, 5–24.

Louca, L., Elby, A., Hammer, D. & Kagey, T. (2004). Epistemological resources: Applying a new epistemological framework to science instruction. *Educational Psychologist*, *39* (1), 57–68.

Mac Callum, R. C., Widaman, K. F., Zhang, S. & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, *4* (1), 87–99.

Maggioni, L., Fox, E. & Alexander, P. A. (2010). An epistemic dimension of competence in the social sciences. *Journal of Social Science Education*, *9* (4), 15–23.

Mansfield, A. F. & Clinchy, B. M. V. (2002). Toward the intetration of objectivity and subjectivity: Epistemological development from 10 to 16. *New Ideas in Psychology*, *20*, 225–262.

Many, J., Howard, F. & Hoge, P. (2002). Epistemology in preservice teacher education: How do beliefs about knowledge affect our students' experiences? *English Education*, *34* (4), 302–322.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57* (3), 519–530.

Marra, R. M. & Palmer, B. (2011). Personal epistemologies and pedagogy in higher education: Did we really mean to say that to our students? In J. Brownlee, G. Schraw & D. Berthelsen (Hrsg.), *Personal epistemology and teacher education* (S. 129–145). New York, London: Routledge.

Marra, R. M., Palmer, B. & Litzinger, T. (2000). The effects of a first-year engineering design course on student intellectual development as measured by the perry scheme. *Journal of Engineering Education*, *89* (1), 39–45.

Marsh, H. W., Balla, J. R. & McDonald, R. D. (1988). Goodness-of-fit indices in confirmtory

factor analysis: The effect of sample size. *Psychological Bulletin*, *103*, 391–410.

Marsh, H. W., Hau, K. T., Balla, J. R. & Grayson, D. (1998). Is more ever too much? the number of indicators per factor in confirmtory factor analysis. *Multivariate Behavioral Research*, *33* (2), 181–220.

Mason, L. (2000). Role of anormalous data and epistemological beliefs in middle school students' theory change about two controversial topics. *European Journal of Psychology of Education*, *15*, 329–346.

Mason, L. (2003). Personal epistemologies and intentional conceptual change. In G. M. Sinatra & P. R. Pintrich (Hrsg.), *Intentional conceptual change* (S. 199–236). Mahwah, N.J.: Erlbaum.

Mason, L. (2016). Psychological perspectives on measuring epistemic cognition. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 375–392). New York and London: Routledge.

Mason, L., Ariasi, N. & Boldrin, A. (2011). Epistemic beliefs in action: Spontaneous reflections about knowledge and knowing during online information searching and their influence on learning. *Learning and Instruction*, *21* (1), 137–151.

Mason, L., Boldrin, A. & Ariasi, N. (2010a). Epistemic metacognition in context: Evaluating and learning online information. *Metacognition and Learning*, *5* (1), 67–90.

Mason, L., Boldrin, A. & Ariasi, N. (2010b). Searching the web to learn about a controversial topic: Are students epistemically active? *Instructional Science*, *38* (6), 607–633.

Mason, L. & Boscolo, P. (2004). Role of epistemological understanding and interest in interpreting a controversy and in topic-specific belief change. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *29* (2), 103–128.

Mason, L., Scirica, F. & Salvi, L. (2006). Effects of beliefs about meaning construction and task instructions on interpretation of narrative text. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *31* (4), 411–437.

McDonald, R. P. (Hrsg.). (1999). *Test theory*. Hillsdale: Erlbaum.

Merk, S. (2016). *Epistemische überzeugungen lehramtsstudierender* (Dissertation). Eberhard Karls Universität Tübingen, Tübingen.

Messick, S. (1989). Validity. In R. L. Linn (Hrsg.), *Educational measurement* (S. 13–104). New York: Macmillan.

Mezirow, J. (Hrsg.). (1991). *Transformative dimensions of adult learning*. San Franciso: Jossey-Bass.

Mezirow, J. (1998). On critical reflection. *Adult Education Quaterly*, *48*, 185–198.

Mokwinski, B. (2010). Die unterschiede in subjektiven theorien zu wissen und wissenserwerb von schüler(inne)n an berufsbildenden schulen und handelslehramtsstudierenden. In S. Behrends, A. Bloemen, B. Mokwinski & W. Schröder (Hrsg.), *Wissen und wissensmanagement.* Oldenburg: BIS.

Mokwinski, B. (2011). *Entwicklungen von epistemologischen überzeugungen in ausgewählten berufsfeldern der dualen berufsausbildung* (Bd. 8). München: Rainer Hampp Verlag.

Moore, W. S. (Hrsg.). (1991). *The perry scheme of intellectual and ethical development: An introduction to the model and major assessment approaches, paper presented at the american*

*educational research association.* Chicago, IL.

Moore, W. S. (1994). Student and faculty epistemology in the college classroom: The perry schema of intellectual and ethical development. In K. W. Pritchard & R. M. Sawyer (Hrsg.), *Handbook of college teaching: Theory and applications* (S. 45–67). Westport, CT: Greenwood.

Moore, W. S. (2002). Understanding learning in a postmodern world: Reconsidering the perrys scheme of intellectual and ethical development. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 17–36). Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Moosbrugger, H. (Hrsg.). (2012). *Testtheorie und fragebogenkonstruktion: Mit 41 tabellen* (2., aktual. u. überarb. Aufl. Aufl.). Berlin: Springer.

Mori, Y. (1999). Epistemological beliefs and language learning beliefs: What do language learners believe about their learning. *Language Learning*, *49* (3), 60–72.

Moschner, B., Gruber, H. & EPI, d. S. (2005). *Epistemologische überzeugungen (epistemic beliefs)* (Bd. Research Report Nr. 18). Regensburg: Universität Regensburg. Lehrstuhl für Lehr-Lern-Forschung.

Moshman, D. (2015). *Epistemic cognition and development: The psychology of justification and truth*. New York: Psychology Press.

Muis, K. R. (2004). Personal epistemology and mathematics: A critical review and synthesis of research. *Review of Educational Research*, *74* (3), 317–377.

Muis, K. R. (2007). The role of epistemic beliefs in self-regulated learning. *Educational Psychologist*, *42* (2), 173–190.

Muis, K. R. (2008). Epistemic profiles and self-regulated learning: Examining relations in the context of mathematics problem solving. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *33* (2), 177–208.

Muis, K. R., Bendixen, L. D. & Haerle, F. C. (2006). Domain-generality and domain-specificity in personal epistemology research: Philosophical and empirical reflections in the development of a theoretical framework. *Educational Psychology Review*, *18* (1), 3–54.

Muis, K. R. & Duffy, M. C. (2013). Epistemic climate and epistemic change: Instruction designed to change students' epistemic beliefs and learning strategies and improve achievement. *Journal of Educational Psychologys*, *105*, 213–225.

Muis, K. R. & Edwards, O. (2009). Examining the stability of achievement goal orientation. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *34* (4), 265–277.

Muis, K. R. & Foy, M. J. (2007). The multilevel effects of teachers' epistemic beliefs on student achievement. In L. D. Bendixen & F. Haerle (Hrsg.), *Personal epistemology in the classroom: Theory, research, and implications for practice.* New York: Cambridge University Press.

Muis, K. R. & Foy, M. J. (2010). The effects of teachers' beliefs on elementary students' beliefs, motivation, and achievement in mathematics. In L. D. Bendixen & F. C. Feucht (Hrsg.), *Personal epistemology in the classroom* (S. 435–469). Cambridge [u.a.]: Cambridge

University Press.

Muis, K. R. & Franco, G. M. (2010). Epistemic profiles and metacognition: Support for the consistency hypothesis. *Metacognition and Learning*, *5* (1), 27–45.

Muis, K. R. & Gierus, B. (2014). Beliefs about knowledge, knowing, and learning: Differences across knowledge types in physics. *Journal of Experimental Education*, *82*, 408–430.

Muis, K. R., Psaradellis, C., Lajoie, S. P., Di Leo, I. & Chevrier, M. (2015). The role of epistemic emotions in mathematics problem solving. *Contemporary Educational Psychology*, *42*, 172–185.

Muis, K. R., Trevors, G. & Chevrier, M. (2016). Epistemic climate for epistemic change. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition.* New York and London: Routledge.

Müller, S. (2009). *Methoden zur erfassung epistemologischer überzeugungen von handelslehramtsstudierenden - eine empirische vergleichsstudie* (Bd. 5). München, Mering: Rainer Hampp.

Müller, S., Heidemann, B. & Rebmann, K. (2006). *Ebid. deutsche übersetzung des epistemic beliefs inventory von schraw, bendixen und dunkle*. Unveröffentlichtes Manuskipt: Universität Oldenburg.

Mundform, D. J., Shaw, D. G. & Ke, T. L. (2005). Minimum sample size recommendations for conducting factor analyses. *Inetrnational Journal of Testing*, *5* (2), 159–168.

Murphy, P. K. (2003). The philosophy in thee: Tracig philosophical influences in educational psychology. *Educational Psychologist*, *38*, 137–145.

Neber, H. & Schommer-Aikins, M. (2002). Self-regulated science learning with highly gifted students: The role of cognitive, motivational, epistemological, and environmental variables. *High Ability Studies*, *13*, 59–74.

Nelson, T. O. & Narens, L. (1994). Why investigate metacognition. In J. Metcalfe & A. Shimamura (Hrsg.), *Metacognition. knowing about knowing* (S. 1–25). Cambridge, Mass: Cambridge University Press.

Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.

Nunnally, J. C. & Bernstein, I. H. (Hrsg.). (1994). *Psychometric theory* (3. Aufl.). New York: McGraw-Hill.

Nussbaum, E. M. & Bendixen, L. D. (2003). Approaching and avoiding arguments: The role of epistemological beliefs, need for cognition, and extraverted personality traits. *Contemporary Educational Psychology*, *28*, 573–595.

Olafson, L. & Schraw, G. (2006). Teachers' beliefs and practices within and across domains. *International Journal of Educational Research*, *45* (1–2), 71–84.

Olafson, L. & Schraw, G. (2010). Beyond epistemology: Assessing teachers' epistemological and ontological worldviews. In L. D. Bendixen & F. C. Feucht (Hrsg.), *Personal epistemology in the classroom* (S. 516–552). Cambridge [u.a.]: Cambridge University Press.

Olsson, U. H., Foss, T., Troye, S. V. & Howell, R. D. (2000). The performance of ml, gls, and wls estimation in structural equation modeling under conditions of misspecification and nonnormality. *Structural Equation Modeling*, *7* (4), 557–595.

Op 't Eynde, P., Corte, E. d. & Verschaffel, L. (2006). Epistemic dimensions of students'

mathematics-related belief systems. *International Journal of Educational Research*, *45* (1–2), 57–70.

Paechter, M., Müller, S., Manhal, S. & Rebmann, K. (Hrsg.). (2007). *Epistemic beliefs and their relation to the application of learning strategies. paper presented at earli.* Budapest.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, *62*, 307–332.

Palmer, B. & Marra, R. M. (2008). Individual domain-specific epistemologies: Implications for educational practice. In M. S. Khine (Hrsg.), *Knowing, knowledge and beliefs* (S. 325–350). Dordrecht: Springer.

Parkinson, M. & Maggioni, L. (2017). The potential of course interventions to change preservice teachers' epistemological beliefs. In G. Schraw, J. Brownlee, L. Olafson & M. VanderVeldt (Hrsg.), *Teachers' personal epistemologies. evolving models for informing practice* (S. 215–238). Charlotte, NC: Information Age Publishing.

Patrick, H. & Pintrich, P. R. (2001). Conceptual change in teachers' intuitive conceptions of learning, motivation, and instruction: The role of motivational and epistemological beliefs. In B. Torf & R. Sternberg (Hrsg.), *Understanding and teaching the intuitive mind* (S. 117–143). Mahwah, NJ: Erlbaum.

Paulsen, M. B. (1998). Domain differences in the epistemological beliefs of college students. *Research in Higher Education*, *39* (4), 365–384.

Paulsen, M. B. & Feldman, K. A. (1999). Student motivation and epistemological beliefs. *New Directions for Teaching and Learning*, *78*, 17–25.

Paulsen, M. B. & Feldman, K. A. (2007). The conditional and interaction effects of epistemological beliefs on the self-regulated learning of college students: Cognitive and behavioral strategies. *Research in Higher Education*, *48* (3), 353–401.

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P. & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The achievement emotions questionnaire (aeq). *Contemporary Educational Psychology*, *36*, 36–48.

Pekrun, R., Goetz, T., Titz, W. & Perry, R. P. (2002a). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, *37*, 91–105.

Pekrun, R., Goetz, T., Titz, W. & Perry, R. P. (2002b). Positive emotions in education. In E. Frydenberg (Hrsg.), *Beyond coping: Meeting goals, visions, and challenges* (S. 149–173). Konstanz: Bibliothek der Universität Konstanz.

Peng, H. & Fitzgerald, G. (2006). Relationsships between teacher education students' epistemological beliefs and their learning outcomes in a case-based hypermedia learning environment. *Journal of Technology and Teacher Education*, *14* (2), 255–285.

Perner, J. & Davies, G. (1991). Understanding the mind as an active information processor: Do young children have a copy theory of mind? *Cognition*, *39* (1), 51–69.

Perry, W. G. (1968). *Patterns of development in thought and values of students in a liberal arts college*. Cambridge, MA: Harvard University Press.

Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart, Winston.

Perry, W. G. (Hrsg.). (1974). *Students as makers of meaning. annual report of the bureau of study counsel*. Harvard University.

Perry, W. G. (1981). Cognitive and ethical growth. In A. Chickering (Hrsg.), *The modern american college* (S. 76–116). San Francisco: Jossey-Bass.

Peter, J. (2015). *Zum wechselseitigen einfluss epistemologischer überzeugungen und förderung von informationskompetenz* (Dissertation). Universität Trier, Trier.

Peters-Burton, E. & Baynard, L. R. (2013). Network analysis of beliefs about the scientific enterprise: A comparison of scientists, middle school science teachers and eighth-grade science students. *International Journal of Science Education*, *35* (16), 2801–2837.

Phan, H. P. (2006). Examination of student learning approaches, reflective thinking and epistemological beliefs: A latent variable approach. *Electronic Journal of Research in Educational Psychology*, *10*, 577–610.

Phan, H. P. (2008). Predicting change in epistemological beliefs, reflective thinking and learning styles: A longitudinal study. *British Journal of Educational Psychology*, *78* (1), 75–93.

Phillips, F. (2001). A research note on accounting students' epistemological beliefs, study strategies, and unstructured problem-solving performance. *Issues in Accounting Education*, *16* (1), 21–39.

Piaget, J. (1950). *The psychology of intelligence*. New York: Routledge.

Piaget, J. (1954). *Das moralische urteil beim kinde*. Zürich: Rascher Orig. 1932.

Piaget, J. (1970). *Psychology and epistemology*. New York: Viking Press.

Piaget, J. (1985). *The equilibration of cognitive structures: The central problem of intellectual development*. Chicago: University of Chicago Press.

Pieschl, S., Stahl, E. & Bromme, R. (2008). Epistemological beliefs and self-regulated learning with hypertext. *Metacognition and Learning*, *3* (1), 17–38.

Pieschl, S., Stahl, E. & Bromme, R. (2013). Adaptation to context as core component of self-regulated learning: The example of complexity and epistemic beliefs. In R. Azevedo & V. Aleven (Hrsg.), *International handbook of metacognition and learning technologies* (S. 53–65). New York: Springer.

Pieschl, S., Stahl, L. & Bromme, R. (2006). *Effects of task difficulty and epistemological beliefs on metacognitive calibration: A pilot-study*. Münster: Universität Münster, ICLS.

Pintrich, P. R. (2002). Future challenges and directions for theory and research on personal epistemology. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 389–414). Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Pintrich, P. R. & Garcia, T. (1993). Intraindividual differences in students' motivation and self-regulated learning. *Zeitschrift für Pädagogische Psychologie*, *7*, 99–107.

Pintrich, P. R., Smith, D. A. F., Garcia, T. & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (mslq)*. NCRIPTAL: University of Michigan: Ann Arbor.

Plöger-Werner, M. (2015). *Epistemologische überzeugungen von erzieherinnen und erziehern: Die bedeutung für das pädagogische handeln in kindertageseinrichtungen*. Wiesbaden: Springer.

Porsch, T. (2011). *Wer kann das wissen? quellenauswahl im kontext der evidenzsuche*. Berlin:

Logos Verlag.

Priemer, B. (2006). Epistemologische überzeugungen von schülern und studenten: Deutschsprachige verfahren der erfassung. In A. Pitton (Hrsg.), *Lehren und lernen mit neuen medien* (S. 120–122). Berlin: LIT-Verlag.

Qian, G. & Alvermann, D. (1995). Role of epistemological beliefs and learned helplessness in secondary school students' learning science concepts from text. *Journal of Educational Psychology*, *87* (2), 282–292.

Qian, G. & Pan, J. (2002). A comparison of epistemological beliefs and learning from science text between american and chinese high school students. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 365–385). Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Raiche, G. & Magis, D. (2010). *nfactors: Parallel analysis and non graphical solutions to the cattell's scree test.* Zugriff auf `https://cran.r-project.org/web/packages/nFactors/nFactors.pdf`

Raiche, G., Riopel, M. & Blais, J.-G. (Hrsg.). (2006). *Non graphical solutions for the cattell's scree test, paper presented at the international annual meeting of the psychometric society.* Montreal, Kanada. Zugriff auf `https://ppw.kuleuven.be/okp/_pdf/Raiche2013NGSFC.pdf`

Raiche, G., Walls, T. A., Magis, D., Riopel, M. & Blais, J.-G. (2013). Non graphical solutions for the cattell's scree test. *Methodology*, *9* (1), 23–29.

Ravindran, B., Greene, B. & DeBacker, T. (2005). Predicting preservice teachers' cognitive engagement with goals and epistemological beliefs. *Journal of Educational Research*, *98* (4), 222–233.

Raykov, T. (2001a). Bias of cronbach's alpha for fixed congeneric measures with correlated errors. *Applied Psychological Measurement*, *25*, 69–76.

Raykov, T. (2001b). Estimation of congeneric scale reliability using covariance structure analysis with nonlinear constraints. *British Journal of Mathematical and Statistical Psychology*, *55*, 145–158.

Redish, E. (1997). *Measuring student expectations in university physics: The mpex survey*. College Park, MD: Physics Education Research Group, University of Maryland.

Reich, K. H., Oser, F. K. & Valentin, P. (1994). Knowing why i now know better: Children's and youth's explanations of their worldview changes. *Journal of Research on Adolescence*, *4* (1), 151–173.

Rest, J. R. (Hrsg.). (1979). *Development in judging moral issues*. University of Minnesota Press: Minneapolis.

Richter, T. (2003). *Epistemologische einschätzungen beim textverstehen*. Lengerich: Pabst.

Richter, T. (2004). *Deutsche übersetzung des fragebogens zur erfassung domänenspezifischer epistemologischer überzeugungen von b. k. hofer* (Unpubliziertes Manuskript). Universität Köln, Köln.

Richter, T. (2011). Cognitive flexibility and epistemic validation in learning from multiple texts. In J. Elen, E. Stahl, R. Bromme & G. Clarebout (Hrsg.), *Links between beliefs and cognitive flexibility* (S. 125–140). Dordrecht: Springer.

Richter, T. & Schmid, S. (2010). Epistemological beliefs and epistemic strategies in self-regulated learning. *Metacognition and Learning*, *5* (1), 47–65.

Rodríguez, L. & Cano, F. (2007). The learning approaches and epistemological beliefs of university students: A cross-sectional and longitudinal study. *Studies in Higher Education*, *32*, 647–667.

Rohrmann, B. (1978). Empirische studien zur entwicklung von antwortskalen für die sozialwissenschaftliche forschung. *Zeitschrift für Sozialpsychologie*, *9*, 222–245.

Rosenberg, S. A., Hammer, D. & Phelan, J. (2006). Multiple epistemological coherences in an eighth-grade discussion of the rock cycle. *Journal of the Learning Sciences*, *15* (2), 261–292.

Rosseel, Y. (2012). lavaan: An r package for structural equation modeling. *Journal of Statistical Software*, *48* (2).

Rost, J. (2004). *Lehrbuch testtheorie - testkonstruktion* (2., vollst. überarb. und erw. Aufl. Aufl.). Bern: Huber.

Royce, J. R. (1959). The search for meaning. *American Scientist*, *47*, 515–535.

Royce, J. R. (1967). Metaphoric knowledge and humanistic psychology. In J. T. F. Bugental (Hrsg.), *Challenges of humanistic psychology* (S. 21–28). New York: McGraw-Hill.

Royce, J. R. (1974). Cognition and knowledge: Psychological epistemology. In E. C. Carterette & M. F. Friedman (Hrsg.), *Handbook of perception. historical and philosophical roots to perception* (S. 149–176). New York: Academic.

Royce, J. R. (1978). Three ways of knowing and the scientific world view. *Methodology and Science*, *11*, 146–164.

Royce, J. R. (1983). Psychological epistemology. *Methodology and Science*, *16*, 164–180.

Royce, J. R. & Mos, L. P. (1980). *Manual: Psycho-epistemological profile*. University of Alberta, Canada: Center for advanced study in theoretical psychology.

Royce, J. R. & Smith, W. A. S. (1964). A note of the development of the psycho-epistemological profile (pep). *Psychological Reports*, *14*, 297–298.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *63*, 581–592.

Rule, D. C. (2003). *The economics of personal epistemology*. University of Nevada, Las Vegas.

Rule, D. C. & Bendixen, L. D. (2010). The integrative model of personal epistemology development: Theoretical underpinnings and implications for education. In L. D. Bendixen & F. C. Feucht (Hrsg.), *Personal epistemology in the classroom* (S. 94–123). Cambridge [u.a.]: Cambridge University Press.

Ruscio, J. & Roche, B. (2012). Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. *Psychological Assessment*, *42* (2), 282–292.

Russell, D. W. (2002). In search of underlying dimension: The use (and abuse) of factor analysis. *Personality and Social Psychological Bulletin*, *28* (12), 1629–1646.

Ryan, M. & Bourke, T. (2013). The teacher as reflexive professional: Making visible the excluded discourse in teacher standards. *Discourse: Studies in the Cultural Politics of Education*, *34*, 411–423.

Ryan, M. P. (1984). Monitoring text comprehension: Individual differences in epistemological standards. *Journal of Educational Psychology*, *76* (2), 248–258.

Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education*, *89*, 634–656.

Sandoval, W. A. (2016). Disciplinary insights into the study of epistemic cognition. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 184–194). New York and London: Routledge.

Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7* (2), 147–177.

Scheele, B. & Schreier, M. (1994). Dialog-konsens-methoden in der empirischen literaturwissenschaft. In A. Barsch, G. Rusch & R. Viehoff (Hrsg.), *Empirische literaturwissenschaft in der diskussion* (S. 278–296). Frankfurt: Suhrkamp.

Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. (2000). Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, *8* (2), 23–74.

Schermelleh-Engel, K. & Wagner, C. (2008). Methoden der reliabilitätsbestimmung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und fragebogenkonstruktion* (S. 114–133). Heidelberg: Springer.

Schiefele, U. & Moschner, B. (1997). *Selbstkonzept, lernmotivation, lernstrategien, epistemologische überzeugungen, instruktionsqualität und studienleistung: Längsschnittliche verläufe und kausale zusammenhänge.*

Schiefele, U., Streblow, L., Ermgassen, U. & Moschner, B. (2003). Lernmotivation und lernstrategien als bedingungen der studienleistung. *Zeitschrift für Pädagogische Psychologie*, *17* (3-4), 185–198.

Schlee, J. (1988). Diagnostik von lernprozessen durch rekonstruktion subjektiver theorien. In H. Eberwein & S. Knauer (Hrsg.), *Handbuch lernprozesse verstehen. wege einer neuen (sonder-)pädagogischen diagnostik* (S. 66–80). Weinheim: Beltz.

Schmid, S. & Lutz, A. (2007). Epistemologische überzeugungen als kohärente laientheorien. *Zeitschrift für Pädagogische Psychologie*, *21* (1), 29–40.

Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, *29* (4), 304–321.

Schoenfeld, A. H. (1983). Beyond the purely cognitive: Belief systems, social cognitions, and metacognition as a driving force in intellectual performance. *Cognitive Science*, *7* (4), 329–363.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. New York: Academic.

Schoenfeld, A. H. (1988). When good teaching leads to bad results: The disasters of well thaught mathematics classes. *Educational Psychologist*, *23*, 145–166.

Schoenfeld, A. H. (1989). Exploration of students' mathematical beliefs and behavior. *Journal for Research in Mathematics Education*, *20*, 338–355.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Hrsg.), *Handbook of research on mathematics, teaching and learning* (S. 334–370). New York: Macmillan.

Schoenfeld, A. H. (1998). Toward a theory of teaching-in-context. *Issues in Education*, *4* (1), 1–94.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, *82* (3), 498–504.

Schommer, M. (1993a). Epistemological development and academic performance among secondary students. *Journal of Educational Psychology*, *85* (3), 406–411.

Schommer, M. (1993b). Comparisons of beliefs about the nature of knowledge and learning among postsecondary students. *Research in Higher Education*, *34*, 355–369.

Schommer, M. (1994). An emerging conceptualization of epistemological beliefs and their role in learning. In R. Garner & P. A. Alexander (Hrsg.), *Beliefs about text and instruction with text* (S. 25–40). Hilldale: Erlbaum.

Schommer, M. (1998). The influence of age and schooling on epistemological beliefs. *Educational Psychology*, *68*, 551–562.

Schommer, M., Calvert, C., Gariglietti, G. & Bajaj, A. (1997). The development of epistemological beliefs among secondary students: A longitudinal study. *Journal of Educational Psychology*, *89* (1), 37–40.

Schommer, M., Crouse, A. & Rhodes, N. (1992). Epistemological beliefs and mathematical text comprehension: Believing it is simple does not make it so. *Journal of Educational Psychology*, *84* (4), 435–443.

Schommer, M. & Dunnell, P. A. (Hrsg.). (2011). *Epistemological beliefs among gifted and non-gifted students, paper presented at the annual meeting of the american educational research association.* San Francisco, USA.

Schommer, M. & Walker, K. (1995). Are epistemological beliefs similar across domains? *Journal of Educational Psychology*, *87* (3), 424–432.

Schommer-Aikins, M. (2002). An evolving theoretical framework of an epistemological belief system. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 103–118). Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Schommer-Aikins, M. (2003). Epistemological world views: A concept that is useful beyond the classroom. *Issues in Education: Contributions from Educational Psychology* (2), 229–232.

Schommer-Aikins, M. (2004). Explaining the epistemological belief system: Introducing the embedded systemic model and coordinated research approach. *Educational Psychologist*, *39* (1), 19–29.

Schommer-Aikins, M. (2008). Applying the theory of an epistemological belief system to the investigation of students' and professors' mathematical beliefs. In M. S. Khine (Hrsg.), *Knowing, knowledge and beliefs* (S. 303–323). Dordrecht: Springer.

Schommer-Aikins, M., Bird, M. & Bakken, L. (2010). Manifestations on an epistemological belief system in preschool to grade twelce classrooms. In L. D. Bendixen & F. C. Feucht (Hrsg.), *Personal epistemology in the classroom* (S. 31–54). Cambridge [u.a.]: Cambridge University Press.

Schommer-Aikins, M., Duell, O. K. & Hutter, R. (2005). Epistemological beliefs, mathematical problem solving beliefs, and academic performance of middle school students.

*Elementary School Journal*, *105*, 289–304.

Schommer-Aikins, M. & Easter, M. (2006). Ways of knowing and epistemological beliefs combined effect on academic performance. *Educational Psychology*, *26*, 411–423.

Schommer-Aikins, M. & Hutter, R. (2002). Epistemological beliefs and thinking about everyday controversial issues. *The Journal of Psychology*, *136*, 5–29.

Schommer-Aikins, M., Mau, W., Brookhart, S. & Hutter, R. (2000). Understanding middle students' beliefs about knowledge and learning using a multidimensional paradigm. *Journal of Educational Research*, *94*, 120–127.

Schraw, G. (2001). Current themes and future directions in epistemological research: A commentary. *Educational Psychology Review*, *13* (4), 451–464.

Schraw, G., Bendixen, L. D. & Dunkle, M. E. (2002). Development and validation of the epistemic belief inventory (ebi). In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 261–275). Mahwah and New Jersey and London: Lawrence Erlbaum Associates.

Schraw, G., Brownlee, J., Olafson, L. & VanderVeldt, M. (2017). *Teachers' personal epistemologies: Evolving models for informing practice*. Charlotte, NC: Information Age Publishing.

Schraw, G., Dunkle, M. E. & Bendixen, L. D. (1995). Cognitive processes in well-defined and ill-defined problem solving. *Applied Cognitive Psychology*, *9*, 523–538.

Schraw, G., Olafson, L. & VanderVeldt, M. (2011). Forstering critical awareness of teachers' epistemological and ontological beliefs. In J. Brownlee, G. Schraw & D. Berthelsen (Hrsg.), *Personal epistemology and teacher education* (S. 149–164). New York, London: Routledge.

Schraw, G., Olafson, L., Weibel, M. & Sewing, D. (2012). Metacognitive knowledge and field-based science learning in an outdoor environmental education program. In A. Zohar & Y. J. Dori (Hrsg.), *Metacognition in science education. trends in current research* (S. 57–77). Dordrecht and Heidelberg and London and New York: Springer.

Schraw, G. & Sinatra, G. M. (2004). Epistemological development and its impact on cognition in academic domains: Reviewer acknowledgment. *Contemporary Educational Psychology*, *29* (2), 95–102.

Schriesheim, C. A., Eisenbach, R. J. & Hill, K. D. (1991). The effect of negation and polar opposite item reversals on questionnaire reliability and validity: An experimental investigation. *Educational and Psychological Measurement*, *51*, 67–78.

Schwab, J. J. (1964). The structure of the disciplines: Meanings and significance. In G. W. Ford & L. Pungo (Hrsg.), *The handbook of educational psychology* (S. 6–30). Chicago: Rand McNally.

Schworm, S. & Gruber, H. (2017). Academic help-seeking: The influence of epistemic beliefs, learning strategies, and goal orientation. In A. Bernholt, H. Gruber & B. Moschner (Hrsg.), *Wissen und lernen. wie epistemische überzeugungen schule, universität und arbeitswelt beeinflussen* (S. 143–162). Münster and New York: Waxmann.

Seidel, T. & Prenzel, M. (2006). Stability of teaching patterns in physics instruction: Findings from a video study. *Learning and Instruction*, *16*, 247–257.

Seidel, T., Schwindt, S., Rimmele, R. & Prenzel, M. (2008). Konstruktivistische überzeugungen von lehrpersonen: Was bedeuten sie für den unterricht. *Zeitschrift für Erziehungswissenschaft*, *10* (9), 259–276.

Seifried, J. (2006). Überzeugungen von (angehenden) handelslehrern. In J. Seifried & J. Abel (Hrsg.), *Empirische lerherbildungsforschung. forschungsstand und perspektiven* (S. 109–127). Münster: Waxmann.

Sinatra, G. M. & Chinn, C. (2012). Thinking and reasoning in science: Promoting epistemic conceptual change. In K. Harris, C. B. McCormick, G. M. Sinatra & J. Sweller (Hrsg.), *Critical theories and models of learning and development relevant to learning and teaching* (Bd. 1, S. 257–282). Washington, D.C.: APA Publications.

Sinatra, G. M. & Kardash, C. M. (2004). Teacher candidates' epistemological beliefs, dispositions, and views on teaching as persuasion. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *29* (4), 483–498.

Sinatra, G. M., Kienhues, D. & Hofer, B. K. (2014). Addressing challenges to public understanding of science: Epistemic cognition, motivated reasoning, and conceptual change. *Educational Psychologist*, *49* (2), 123–138.

Sinatra, G. M. & Pintrich, P. R. (Hrsg.). (2003). *Intentional conceptual change: Includes bibliographical references and indexes*. Mahwah, N.J.: Erlbaum.

Smith, C. L., Maclin, D., Houghton, C. & Hennessey, M. G. (2000). Sixth-grade students' epistemologies of science: The impact of school science experiences on epistemological development. *Cognition and Instruction*, *18* (3), 349-–422.

Sodian, B. & Kristen, S. (2016). Theory of mind. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 68–85). New York and London: Routledge.

Sosu, E. M. & Gray, D. S. (2012). Investigating change in epistemic beliefs: An evaluation of the impact of student teachers' beliefs on instructional preference and teaching competence. *International Journal of Educational Research*, *53*, 80–92.

Spiro, R. J., Coulson, R. L., Feltovich, P. J. & Anderson, D. K. (1994). Cognitive flexibility theory: Advance knowledge acquisition in ill-structured domains. In R. B. Ruddell, M. R. Ruddell & H. Singer (Hrsg.), *Theoretical models and processes of reading* (S. 602–615). Newark: International Reading Association.

Spiro, R. J., Feltovich, P. J. & Coulson, R. L. (1996). Two epistemic world-views: Prefigurative schemas and learning in complex domains. *Applied Cognitive Psychology*, *10*, 51–61.

Spiro, R. J. & Jehng, J. C. (1990). Cognitive flexibility and hypertext: Theory and technology for the nonlinear and multidimensional traversal of complex subject matter. In D. Nix & R. J. Spiro (Hrsg.), *Cognition, education, and multimedia: Explorations in high technology* (S. 163–205). Hillsdale, NJ: Erlbaum.

Stahl, E. & Bromme, R. (2007). The caeb: An instrument for measuring connotative aspects of epistemological beliefs. *Learning and Instruction*, *17* (6), 773–785.

Stahl, E., Pieschl, S. & Bromme, R. (2006). Task complexity, epistemological beliefs and metacognitive calibration: An exploratory study. *Journal of Educational Computing Research*, *35* (4), 319–338.

Stallmann, F. (2007). *The influence of epistemological beliefs and academic self-concept on metacognitive calibration in german 12th graders* (Unpublizierte Diplomarbeit). Universität Müenster, Müenster.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, *25* (2), 173–180.

Stevens, R., Wineburg, S., Herrenkohl, L. R. & Bell, P. (2005). Comparative understanding of school subjects: Past, present, and future. *Review of Educational Research*, *75* (2), 125–127.

Stewart, I. (1987). *The problem of mathematics*. New York: Oxford University Press.

Stodolsky, S. S. (1985). Telling math: Origins of math aversion and anxiety. *Educational Psychologist*, *20*, 125–133.

Stodolsky, S. S. (1988). *The subject matters: Classroom activities in math and social studies*. University of Chicago Press: Chicago.

Stodolsky, S. S., Salk, S. & Glaessner, B. (1991). Students views about learning math and social studies. *American Educational Research Journal*, *28*, 89–116.

Streblow, L. & Schiefele, U. (2006). Lernstrategien im studium. In H. Mandl & H. F. Friedrich (Hrsg.), *Handbuch lernstrategien* (S. 352–364). Göttingen: Hogrefe.

Strømsø, H. I. & Bråten, I. (2011a). Personal epistemology in higher education: Teachers' beliefs and the role of faculty training programs. In J. Brownlee, G. Schraw & D. Berthelsen (Hrsg.), *Personal epistemology and teacher education* (S. 54–67). New York, London: Routledge.

Strømsø, H. I. & Bråten, I. (2011b). Personal epistemology in higher education: Teachers' beliefs and the role of faculty training programs. In J. Brownlee, G. Schraw & D. Berthelsen (Hrsg.), *Personal epistemology and teacher education* (S. 54–67). New York, London: Routledge.

Strømsø, H. I. & Kammerer, Y. (2016). Epistemic cognition and reading for understanding in the internet age. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 230–246). New York and London: Routledge.

Tabachnick, B. G. & Fidell, L. S. (2001). *Using multivariate statistics* (Bd. 4). Boston, MA: Allyn & Bacon.

Tabak, I. & Weinstock, M. (2008). A sociocultural exploration of epistemological beliefs. In M. S. Khine (Hrsg.), *Knowing, knowledge and beliefs* (S. 177–195). Dordrecht: Springer.

Tabak, I. & Weinstock, M. (2011). If there is no one right answer? the epistemological implications of classroom interactions. In J. Brownlee, G. Schraw & D. Berthelsen (Hrsg.), *Personal epistemology and teacher education* (S. 180–194). New York, London: Routledge.

Team, R. C. (2017). *R: A language and environment for statistical computing*. Wien, Österreich. Zugriff auf `http://www.r-project.org`

Thoermer, C. & Sodian, B. (2002). Science undergraduates' and graduates' epistemologies of science: The notion of interpretice frameworks. *New Ideas in Psychology*, *20*, 263–283.

Thurstone, L. L. & Chave, E. J. (1929). *The measurement of attitude: A psychophysical method and some experiments with a scale for measuring attitude toward the church*. Chicago, Illinois: University of Chicago Press.

Tillma, H. (2011). Looking into mirrors: Teacher educators' dilemmas in constructing pedagogical understanding about their teaching. In J. Brownlee, G. Schraw & D. Berthelsen (Hrsg.), *Personal epistemology and teacher education* (S. 40–53). New York, London: Routledge.

Trautwein, U., Lüdke, O. & Beyer, B. (2004). Rauchen ist tödlich, computerspiele machen aggressiv? allgemeine und theorienspezifische epistemologische überzeugungen bei studierenden unterschiedlicher fachrichtungen. *Zeitschrift für Pädagogische Psychologie*, *18* (3/4), 187–199.

Trautwein, U. & Lüdtke, O. (2004). Wissenschaftspropädeutik in der gymnasialen oberstufe. In O. Köller, R. Watermann, U. Trautwein & O. Lüdtke (Hrsg.), *Wege zur hochschulreife in baden-württemberg: Tosca - eine untersuchung an allgemein bildenden und beruflichen gymnasien: Literaturangaben* (S. 351–366). Opladen: Leske und Budrich.

Trautwein, U. & Lüdtke, O. (2007). Epistemological beliefs, school achievement, and college major: A large-scale longitudinal study on the impact of certainty beliefs. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *32* (3), 348–366.

Tsai, C.-C. (1998). An analysis of scientific epistemological beliefs and learning orientations of taiwanese eighth graders. *Science Education*, *82*, 473–489.

Tsai, C.-C. (1999). Laboratory exercises help me memorize the scientific truths: A study of eighth graders' scientific epistemological views and learning in laboratory activities. *Science Education*, *83* (6), 654–674.

Tsai, C.-C. (2002). Nested epistemologies: Science teachers' beliefs of teaching, learning and science. *International Journal of Science Education*, *24*, 771–783.

Tsai, C.-C. (2006). Re-interpreting and reconstructing science: Teachers' view changes towards the nature of science by courses of science education. *Teaching and Teacher Education*, *22*, 363–375.

Urhahne, D. & Harms, U. (2006). Instruktionale unterstützung beim lernen mit computersimulationen. unterrichtswissenschaft. *Zeitschrift für Pädagogische Psychologie*, *34* (4), 258–377.

Urhahne, D. & Hopf, M. (2004). Epistemologische überzeugungen in den naturwissenschaften und ihre zusammenhänge mit motivation, selbstkonzept und lernstrategien. *Zeitschrift für Didaktik der Naturwissenschaften*, *10*, 71–87.

Urhahne, D., Kremer, K. & Mayer, J. (2008). Welches verständnis haben jugendliche von der natur der naturwissenschaften? entwicklung und erste schritte zur validierung eines fragebogens. *Unterrichtswissenschaft*, *36*, 71–93.

Valanides, N. & Angeli, C. (2005). Effects of instruction on changes in epistemological beliefs. *Epistemological Development and Its Impact on Cognition in Academic Domains*, *30* (3), 314–330.

Valanides, N. & Angeli, C. (2011). Personal epistemology and ill-defined problem solving in solo and dyadic contexts. In J. Brownlee, G. Schraw & D. Berthelsen (Hrsg.), *Personal epistemology and teacher education* (S. 195–209). New York and London: Routledge.

VanSledright, B. & Maggioni, L. (2016). Epistemic cognition in history. In J. A. Greene, W. A. Sandoval & I. Bråten (Hrsg.), *Handbook of epistemic cognition* (S. 128–146). New

York and London: Routledge.

Veenman, M. V. J. (2012). Metacognition in science education: Definitions, constituents, and their intricate relation with cognition. In A. Zohar & Y. J. Dori (Hrsg.), *Metacognition in science education. trends in current research* (S. 21–36). Dordrecht and Heidelberg and London and New York: Springer.

Veenman, M. V. J., van Hout-Wolters, B. H. A. M. & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, *1* (1), 3–14.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika*, *41*, 321–327.

Velicer, W. F., Eaton, C. A. & Fava, J. L. (2000). Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In R. D. Goffin & E. Helmes (Hrsg.), *Problems and solutions in human assessment: Honoring douglas n. jackson at seventy* (S. 41–71). Boston: Kluwer.

Velicer, W. F. & Fava, J. L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods*, *3* (2), 231–251.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, *4* (1), 45–69.

Wahl, D. (1979). Methodische probleme bei der erfassung handlungsleitender und handlungsrechtfertigender subjektiver psychologischer theorien von lehrern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *11* (3), 208–217.

Walker, S., Brownlee, J., Exley, B., Woods, A. & Whiteford, C. (2011). Personal epistemology in preservice teachers: Belief changes throughout a teacher education course. In J. Brownlee, G. Schraw & D. Berthelsen (Hrsg.), *Personal epistemology and teacher education* (S. 84–99). New York, London: Routledge.

Walker, S., Brownlee, J., Whiteford, C., Exley, B. & Woods, A. (2012). A longitudinal study of change in preservice teachers' personal epistemologies. *Australian Journal of Teacher Education*, *37* (5), 24–35.

Walton, M. D. (2000). Say it's a lie or i'll punch you: Naive epistemology in classroom conflict episodes. *Discourse Processes*, *29* (2), 113–136.

Weinstein, C. E. & Mayer, R. E. (1986). The teaching of learning strategies. In M. C. Wittrock (Hrsg.), *Handbook of research on teaching* (S. 315–327). New York: Macmillan.

Weinstock, M. & Cronin, M. A. (2003). The everyday production of knowledge: Individual differences in epistemological understanding and juror-reasoning skill. *Applied Cognitive Psychology*, *17*, 161–181.

Weinstock, M. & Roth, G. (2011). Teachers' personal epistemologies as predictors of support for their students' autonomy. In J. Brownlee, G. Schraw & D. Berthelsen (Hrsg.), *Personal epistemology and teacher education* (S. 165–179). New York, London: Routledge.

Weinstock, M. & Zviling-Beiser, H. (2009). Separating academic and social experience as potential factors in epistemological development. *Learning and Instruction*, *19* (3), 287–298.

Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.

Wellman, H. M., Cross, D. & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child development*, *72*, 655–684.

Wellman, H. M. & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology*, *43* (1), 337–375.

West, S. G., Finch, J. F. & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Hrsg.), *Structural equation modeling: Concepts, issues, and applications* (S. 56–75). Thousand Oaks, CA: Sage Publications.

Whitmire, E. (2004). The relationship between undergraduates' epistemological beliefs, reflective judgment, and their information-seeking behavior. *Information Processing and Management*, *40* (1), 97–111.

Wigfield, A. & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, *25*, 68–81.

Wild, K.-P. (2000). *Lernstrategien im studium*. München: Waxmann.

Wild, K.-P. (2005). Individuelle lernstrategien von studierenden: Konsequenzen für die hochschuldidaktik und die hochschullehre. *Beiträge zur Lehrerbildung*, *23* (2), 191–206.

Wild, K.-P. (2011). *List-r: Inventar zur erfassung von lernstrategien im studium (revidiert)*. Regensburg: Universität Regensburg.

Wild, K.-P. & Schiefele, U. (1993). Induktiv versus deduktiv entwickelte fragebogenverfahren zur erfassung von merkmalen des lernverhaltens. *Unterrichtswissenschaft*, *21*, 312–326.

Wild, K.-P. & Schiefele, U. (1994). Lernstrategien im studium. ergebnisse zur faktorenstruktur und reliabilität eines neuen fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *15*, 185–200.

Wildenger, L., Hofer, B. K. & Burr, J. E. (2010). Epistemological development in very young knowers. In L. D. Bendixen & F. C. Feucht (Hrsg.), *Personal epistemology in the classroom: Theory, research, and implications for practice* (S. 220–257). Cambridge, UK: Cambridge University Press.

Windschitl, M. & Andre, T. (1998). Using computer simulations to enhance conceptual: The roles of constructivist instruction and student epistemological beliefs. *J. Res. Sci. Teaching*, *23* (2), 145–160.

Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, *83* (4), 73–87.

Wineburg, S. S. (1996). The psychology of learning and teaching history. In D. C. Berliner & R. C. Calfee (Hrsg.), *The handbook of educational psychology* (S. 423–437). New York: Macmillan.

Winne, P. H. & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Hrsg.), *Metacognition in educational theory and practice* (S. 277–304). Mahwah, NJ: Lawrence Erlbaum Associates.

Wood, P. & Kardash, C. (2002). Critical elements in the design and analysis of studies of epistemology. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal epistemology* (S. 231–260).

Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Yadav, A., Herron, M. & Samarapungavan, A. (2011). Personal epistemology in preservice teacher education. In J. Brownlee, G. Schraw & D. Berthelsen (Hrsg.), *Personal epistemology and teacher education* (S. 25–39). New York, London: Routledge.

Yadav, A. & Koehler, M. (2007). The role of epistemological beliefs in preservice teachers' interpretation of video cases of early-grade literacy instruction. *Journal of Technology an Teacher Education*, *15* (3), 335–361.

Yang, Y.-F. & Tsai, C.-C. (2010). Conceptions of and approaches to learning through online peer assessment. *Learning and Instruction*, *20* (1), 72–83.

Youn, I. (2000). The cultural specificity of epistemological beliefs about learning. *Asian Journal of Social Psychology*, *3*, 87–105.

Youn, I., Yang, K. & Choi, I. (2001). An analysis ot the nature of epistemological beliefs: Integrating factors affecting the epistemological development of south korean high school students. *Asia Pacific Education Review*, *99* (3), 10–21.

Zinn, B. (2013). *Überzeugungen zu wissen und wissenserwerb: Empirische untersuchungen zu den epistemologischen überzeugungen lernender*. Münster and New York and München and Berlin: Waxmann.

Zwick, W. R. & Velicer, W. F. (1986). Comarison of five rules for determining the number of components to retain. *Psychological Bulletin*, *99* (3), 432–442.