

Cross-curricular problem solving

Entwicklung und Erprobung eines Testverfahrens und
Analyse von Bedingungen fächerübergreifender
Problemlösekompetenzen bei Studierenden

Inaugural-Dissertation zur Erlangung des Doktorgrades
der Pädagogisch-Philosophischen Fakultät

der

Katholischen Universität Eichstätt-Ingolstadt

vorgelegt von

Nicole Romana Heigl, Johannisstrasse 7, 54290 Trier

2014

Referent: Prof. Dr. Joachim Thomas

Koreferent: Prof. Dr. Ulrich Bartosch

Tag der Disputation: 03.07.2014

Zusammenfassung

Probleme in neuen Situationen zu lösen prägt unseren Alltag und Beruf. Die Fähigkeit, sich mit komplexen und wechselnden Sachverhalten auseinandersetzen zu können ist dementsprechend wichtig, um sowohl an der Gesellschaft als auch am beruflichen Leben erfolgreich teilnehmen zu können und sich neues Wissen zu erschließen. Aktuelle Studien zeigen, dass die Anwendung von im Studium erworbenem Wissen und Können von Arbeitgebern und Studierenden nicht als ausreichend eingeschätzt wird, um diese Anforderungen bewältigen zu können. Zurückgeführt wird dies auch darauf, dass bislang zu wenig sogenannte berufsbefähigende Kompetenzen im Studium vermittelt wurden. Um diese Lücke zu schließen, wurden in den letzten Jahren unterschiedliche Projekte angestoßen: So begann die OECD vor einigen Jahren mit der Analyse derjenigen Kompetenzen, die einen mündigen Bürger auszeichnen und definierte sogenannte *life skills*, bzw. *key competencies*. Im Rahmen der PISA-Studien wurden auf schulischer Ebene auch Fähigkeiten evaluiert, neue Probleme lösen zu können. Auf Hochschulebene wurde nun das Hochschulstudium reformiert. Mit der Etablierung von Bachelor- und Masterstudiengängen sollte der Fokus auch auf die Berufsbefähigung der Studierenden gelegt werden. Problematisch in diesem Zusammenhang ist, dass bislang ungeklärt ist, wie diese berufsbefähigenden Kompetenzen, als Teil sogenannter *academic skills*, das heißt, an der Hochschule zu erwerbender Fähigkeiten, definiert und gemessen werden können und inwieweit Studierende beim Erwerb dieser Kompetenzen unterstützt werden müssen, bzw. können.

In der vorliegenden Studie wurde ein theoretisches Modell der Kompetenz fächerübergreifenden Problemlösens – das heißt, Probleme in neuen Situationen lösen zu können – und ein diagnostisches Instrument zur Evaluation dieser Kompetenz entwickelt und getestet. Dabei wurden auch Bedingungsfaktoren fächerübergreifenden Problemlösens identifiziert. Anhand der empirischen Überprüfung des Modells konnten bedeutsame Modellmerkmale identifiziert werden.

Die Befunde weisen darauf hin, dass insbesondere diejenigen Probleme schwierig zu lösen sind, die eine Wahrscheinlichkeitsschätzung beinhalten oder es erfordern, komplexe Informationen zu ordnen. Dabei spielt auch die Modalität der Problemelemente und die Reflektiertheit der Studierenden im Umgang mit Wissen (epistem. *beliefs*), sowie die Einschätzung der Wirksamkeit des Handelns der Studierenden eine Rolle. Insgesamt zeigen die Befunde, dass sich die Studierenden verschiedener Semester und Studiengänge in Abhängigkeit der Anforderung des jeweiligen Problems unterscheiden. Dies lässt den Schluss zu, dass sich Problemlösefähigkeiten im Laufe des Studiums verändern. Weiterhin lässt dies auch die Frage zu, inwieweit Studiengänge unterschiedlich auf fächerübergreifendes Problemlösen vorbereiten.

Die Befunde verdeutlichen die Notwendigkeit, Studierenden zu vermitteln, inwieweit fachspezifische Fähigkeiten auch fachübergreifend angewendet werden können, sowie Selbstwirksamkeit und epistemologische *beliefs* bei der Lehre zu berücksichtigen. Die vorliegende Studie geht über bestehende Studien zum Problemlösen und fächerübergreifenden Problemlösen hinaus, da Faktoren des Problems und der Person gemeinsam betrachtet und Ableitungen für die Lehre mit bedacht wurden.

Abstract

In our every-day as well as in our professional life, we are used to solving problems in new situations. The ability to deal with complex and changing circumstances is important in order to succeed in professional as well as social life. However, current studies show that both employers and employees often feel that knowledge and skills learned in academic studies are not sufficient to deal with professional problems. This is often attributed to a lack of teaching so-called *berufsbefähigende Kompetenzen* (job-relevant competencies) in academic studies. In order to ensure a better successful professional practice in the future, different projects have been introduced in the past few years: OECD began the analysis of competencies that characterize an informed citizen, and defined so called *life skills* and *key competencies*. The subsequent PISA-studies of European primary and secondary education assessed, among others, competencies for dealing with new problems. Also, the European Higher Education System was being reformed. In the newly established Bachelor and Master programs, the *employability of graduates* was supposed to be a key factor. However, it is still unclear how job-relevant competencies, as part of so called academic skills (that is, skills learned in a higher education program), may be defined and evaluated. It also remains unclear how these competencies could and should be taught to students.

This study develops and tests a theoretical model and a diagnostic tool for analyzing *cross-curricular problem solving competencies* – that is, the ability to solve problems in new situations. For this, conditions for cross-curricular problem-solving were identified. With the empirical test the study was able to identify significant model features.

The results show that especially those problems were hard to solve that involved assessments of probability or required organizing complex information. Important factors were the type of problems involved, the degree to which students were aware of their knowledge (epistemological beliefs), as well as their assessment of the effectiveness of their actions. The results show that there are differences between students of different fields and academic years with regard to certain types of problem. That is, the ability to solve problems changes over the years of study. This raises the question, whether some disciplines prepare their students better for cross-curricular problem-solving than others.

The results show the necessity to teach students how subject-specific knowledge can be applied to other areas, and to consider self-efficacy and epistemological beliefs in teaching. By examining both personal and problem-related factors, and by presenting implications for the improvement of teaching in higher education, this study goes beyond previous studies on problem-solving and cross-curricular problem-solving.

Inhaltsverzeichnis

Einleitung : Entwicklung fächerübergreifender Kompetenzen als Schlüssel zum Erfolg	8
Teil I: Herausforderungen im Zusammenhang mit der Messung fächerübergreifender Problemlösekompetenzen	12
1. <i>Academic skills</i> und fächerübergreifendes Problemlösen – Zur Ausgangslage der bildungs- politischen Diskussion und Einordnung des Untersuchungsgegenstandes	12
1.1 Fragestellungen im Zusammenhang mit der Diskussion um <i>academic skills</i>	13
1.2 Ansatzpunkte zur Operationalisierung von <i>academic skills</i> und Einordnung fächer- übergreifender Problemlösekompetenzen	16
2. Der Kompetenzbegriff: Konjunktur und Problematik	19
2.1 Systematisierungsansätze	20
2.2 Das funktionale Kompetenzkonzept der Psychologie	22
3. Kompetenz als Bewältigung von Handlungsanforderungen und Konsequenzen für die Messung des Kompetenzkonstrukts	23
3.1 Eine pragmatische Arbeitsdefinition	23
3.2 Anforderung an die Kompetenzmessung	25
4. Zusammenfassung	28
Teil II: Theoretische Verortung fächerübergreifender Problemlösekompetenzen	29
5. Fächerübergreifendes Problemlösen: Theorien, Positionen, Befunde	29
5.1. Problemlösen – Theorien und Modelle	31
<i>Zentrale Fragestellungen im Zusammenhang mit Problemlösen</i>	31
<i>Antwortmodelle: Theorien des Problemlösens</i>	32
<i>Zusammenfassung</i>	35
5.2 Kognitive Anforderungen beim Problemlösen	35
<i>Problemlösen als kognitive Verarbeitung von Aussagen</i>	36
<i>Fehlerquellen im logischen Schließen</i>	38
<i>Zusammenfassung</i>	43
5.3 Kontextgebundenheit fächerübergreifender Problemlösekompetenzen	44
<i>Fächerübergreifendes Problemlösen unter dem Gesichtspunkt analogen Problemlösens und Transfers</i>	45
<i>Systematisierungsansätze</i>	47
<i>Fächer(un)abhängigkeit fächerübergreifender Problemlösekompetenzen</i>	49
5.4 Modellierung fächerübergreifenden Problemlösens in Abhängigkeit des Problems	51
<i>Wie nimmt das Problem Einfluss auf seine Lösung?</i>	51
<i>Fächerübergreifendes Problemlösen: Eine Frage des Problems?</i>	52

5.5 Fächerübergreifendes Problemlösen: Eine differentialpsychologische Sicht auf das Konstrukt	55
<i>Fächerübergreifendes Problemlösen und Intelligenz</i>	56
<i>Die Rolle von Vorwissen und Expertise</i>	58
<i>Problemlösen, affektive und metakognitive Faktoren</i>	60
<i>Alters- und Geschlechtseffekte</i>	66
5.6 Die Frage der Messung fächerübergreifender Problemlösekompetenzen	68
<i>Anforderungen an die Messung fächerübergreifender Problemlösekompetenzen</i>	68
<i>Überblick über diagnostische Instrumentarien zur Kompetenzmessung</i>	69
6. Zusammenfassung	73
7. Forschungsfragen und Hypothesen	75
Teil III: Testinstrument und methodisches Design der Studie	79
8. Anforderungen an das Testverfahren	80
8.1 Ziele des Testinstrumentes	80
8.2 Zielstichprobe des Tests	80
8.3 Formale Anforderungen an die Konstruktion des Verfahrens	80
8.4 Darstellung der begleitenden Information	81
8.5 Anforderungen an Testformat und Itemdesign	83
8.6 Reihenfolgeeffekte und Itemfolge	84
8.7 Abgrenzung des vorliegenden Testverfahrens	85
9. Beschreibung des Testverfahrens: Das formale Testdesign	85
9.1 Die formale Test- und Itemstruktur	86
9.2 Operationalisierung fächerübergreifender Problemlösekompetenzen: Test- und Itemdesign	88
<i>Testszenario und Itembeispiele</i>	90
<i>Testmodell</i>	99
9.3 Ergebnisraum und Auswertungsschema	100
9.4 Umgang mit fehlenden Werten	100
10. Pilotierung des Testinstruments	102
10.1 Stichprobe der Vorstudie	102
10.2 Durchführung der Vorstudie	102
10.3 Ergebnisse der Vorstudie	103
<i>Kennwerte des Subtests Situationsanalyse</i>	104
<i>Kennwerte des Subtests Zielevaulation</i>	106
<i>Kennwerte des Subtests Handlungsplanung</i>	107
10.4 Umgang mit fehlenden Werten	108

10.5 Diskussion der Güte des Testverfahrens	109
<i>Objektivität</i>	109
<i>Reliabilität</i>	109
<i>Validität</i>	110
11. Beschreibung der Hauptstudie	111
11.1 Tabellarische Übersicht über Testaufgaben und Items	111
11.2 Untersuchungsdesign	111
11.3 Stichprobe	112
11.4 Untersuchungsverlauf	112
11.5 Untersuchungszeitraum	113
11.6 Varianz der Hauptstichprobe	114
<i>Studierende, Studiengänge und Semesterverhältnis</i>	114
<i>Geschlechterverhältnis</i>	115
<i>Studierende: Alter, Bildungserfahrung und Bildungshintergrund</i>	115
11.7 Beschreibung der weiteren Verfahren	116
<i>Studieninteresse</i>	116
<i>Selbstwirksamkeit</i>	117
<i>Epistemologische Überzeugungen</i>	117
<i>Intelligenzscreening</i>	119
Teil IV: Befunde der Hauptstudie	120
12. Verteilung der Daten	120
13. Fragenkomplex I: Konstruktklärung	120
13.1 Verwendete Verfahren	121
13.2 Binnenstruktur des Konstrukts	123
13.3 Modellüberprüfung	124
13.4 Abgrenzung der Testintelligenz	125
13.5 Domänenspezifität fächerübergreifenden Problemlösens	126
14. Fragenkomplex II: Wirkungszusammenhänge zwischen personbezogenen Variablen und fächerübergreifenden Problemlösekompetenzen	128
14.1 Verwendete Verfahren	128
14.2 Zusammenhänge zwischen personalen Variablen	129
14.3 Differentialpsychologische Faktoren als Kovariaten der Problemlöseleistung	132
14.4 Testintelligenz und Expertise	133
14.5 Zusammenfassung der Wechselwirkungen und Skizze eines Strukturmodells	134
15. Fragenkomplex III: Fächerübergreifende Problemlösekompetenzen bei Studierenden	138
15.1 Verwendete Verfahren	139
15.2 Fächerübergreifendes Problemlösen – eine Frage der Expertise?	139

<i>Der Einfluss des Semesters</i>	140
... und die Rolle epistemologischer Überzeugungen	143
<i>Der Einfluss der Studiendisziplin</i>	144
... und die Rolle epistemologischer Überzeugungen	145
15.3 Fehleranalyse mit Blick auf Zielevaluation und Handlungsplanung	146
<i>Muster bei fehlenden Werten</i>	146
<i>Schwierigkeiten bei Testitems zur Zielevaluation</i>	146
<i>Schwierigkeiten bei Testitems zur Handlungsplanung</i>	148
<i>Die Schwierigkeit der korrekten Platzierung</i>	149
15.4 Zusammenhänge zwischen Teilkompetenzen fächerübergreifenden Problemlösens	150
15.5 Gesamtmodell zur Vorhersage fächerübergreifender Problemlösekompetenz	153
Teil V: Diskussion und Ausblick – Einordnung der zentralen Befunde	157
16. Das Testverfahren zur fächerübergreifenden Problemlösekompetenz: Theoretisches Konzept, Güte des Testverfahrens und Binnenstruktur des Konstrukts	157
17. Die Frage der Domänenspezifität fächerübergreifender Problemlösekompetenzen	159
18. Fächerübergreifende Problemlösekompetenzen bei Studierenden	161
18.1 Fächerübergreifende Problemlösekompetenzen bei Studierenden und die Rolle der Expertise	161
18.2 Weitere Schwierigkeiten beim fächerübergreifenden Problemlösen	163
18.3 Die Rolle subjektiver Überzeugungen (I): Epistemologische Überzeugungen und Selbstwirksamkeit	164
18.4 Die Rolle subjektiver Überzeugungen (II): Transfer	167
18.5 Studieninteresse	168
18.6 Kompetenzstrukturen	169
19. Schlussfolgerungen für die Lehre und Ableitungen für die Analyse von <i>academic skills</i>	170
20. Einschränkungen und Grenzen	172
20.1 Das Testinstrument	172
20.2 Prüfung der konvergenten und diskriminanten Validität	173
20.3 Methode und Stichprobe	174
20.4 Kovariaten fächerübergreifenden Problemlösens	174
20.5 Grenzen und Einschränkungen im Hinblick auf Schlussfolgerungen für die Lehre	176
21. Zusammenfassung und Ausblick	176
Literaturverzeichnis	179
Anhang	220
Testverfahren	
Erklärung	

Einleitung

“[...] [B]eing able to solve problems is thought to be a characteristic of an ‘educated’ person.”

Voss, 1989, S. 251

Entwicklung fächerübergreifender Kompetenzen als Schlüssel zum Erfolg

Problemlösen wird als eine wichtige kognitive Fähigkeit sowohl in Alltag als auch Beruf angesehen (u.a. Fleischer, Wirth, Rumann & Leutner, 2010; Strub, Dubach & Oesch, 2008; Manninen & Hobrough, 2000; Austin et. al., 1997; Klemp & McClelland, 1986). Sie stellt ein bedeutendes Bildungsziel dar, da sie es einerseits ermöglicht, sich neues Wissen zu erschließen und andererseits eine zentrale berufliche Anforderung darstellt (Baumert, Klieme, Neubrand, Prenzel, Schiefele et. al., o.J.; Fleischer et. al., 2010; Mayer, 2010). Problemlösen wird dabei jedoch nicht nur als fachspezifische, sondern auch als fächerübergreifende Kompetenz begriffen, das heißt, das zu lösende Problem lässt sich nicht unbedingt immer einer Fachdomäne zuordnen. Zum Lösen dieser Probleme ist dann weniger Fachwissen, sondern vielmehr der Transfer fachspezifisch erworbener Problemlösekompetenzen notwendig (OECD, 2004). Für das spätere Berufsleben, mit seinen wechselnden Anforderungen, der Notwendigkeit des Umgangs mit unbekanntem Situationen und der erforderlichen Anpassung, wird fächerübergreifendem Problemlösen deshalb eine zentrale Rolle zugeschrieben (OECD, 2004; Billing, 2000):

“Problem solving and reasoning are frequently in the spotlight of educational reform. [...] [S]tudents must be able to use what they have learned to solve new problems. [...] A major goal of education is to help students to become effective problem solvers [...]” (Mayer, 2010, S. 273; auch Bennett, Dunne, Carré, 2000; De la Harpe et. al., 2000; Barnett, 1994)

Das Konzept fächerübergreifender Problemlösekompetenzen ist somit ein Schlüsselthema, es soll im Studium die fachspezifische Perspektive ergänzen und eine Brücke zwischen beruflicher Anforderung und Berufsvorbereitung schlagen (Rychen & Salganik, 2001; Barnett, 2000a, 2000b, 2000c; Bennett, Dunne & Carré, 2000; Bridges, 2000; Mertens, 1974). Und obwohl die Bedeutung fächerübergreifender Kompetenzen schon seit langem bekannt ist (Mertens, 1974), gewann die Diskussion erst wieder im Zusammenhang mit dem Bologna-Prozess an Dynamik.

Natürlich werden im Unterricht neben fachlichen Kompetenzen in klar definierten curricularen Leistungsbereichen immer auch unweigerlich fachübergreifende Kompetenzen gelehrt und gelernt (Peschar, 2004; Klieme, Artelt & Stanat, 2001). Auch in fachspezifischen Leistungsmessungen werden fachübergreifende Kompetenzen implizit mit erfasst. Allerdings konzentrierten sich Lehre und

Forschung, insbesondere die Frage der Kompetenzmessung bisher überwiegend auf Kompetenzen in Fachdisziplinen. Eine Ausnahme sind Projekte wie bspw. *ProfilPASS* (DIE), Kompetenznachweis Lernen im sozialen Umfeld (DJI) oder der *Kompetenznachweis Kultur*, die im schulischen Bereich oder im Bereich der beruflichen Neuorientierung angesiedelt sind (bspw. Thomas, 2009). Der Fokus auf fachübergreifende Kompetenzen im Hochschulbereich stellt nun eine notwendige Ergänzung der funktional definierten und auf bestimmte Leistungsbereiche bezogenen Kompetenzmessungen dar. Er verweist auf die Herausforderung, auch fächerübergreifende Kompetenzen messbar und bewertbar zu machen. Eine systematisch geleitete und empirisch ausgerichtete Beschäftigung mit dem Konstrukt ist dementsprechend unerlässlich.

Im Zusammenhang damit erweisen sich allerdings verschiedene Aspekte als problematisch. Zum einen ist bis heute unklar, wie das Konstrukt theoretisch gefasst werden kann, und zwar sowohl bezogen auf fächerübergreifende Kompetenzen im Allgemeinen als auch fächerübergreifende Problemlösekompetenzen im Spezifischen (Baumert et. al., 2000; auch Blömeke, Zlatkin-Troitschanskaia, Kuhn & Fege, 2013). Was unter dem Konstrukt genau zu verstehen ist, wie es theoretisch und empirisch systematisiert werden kann, bleibt trotz der Fülle an Literatur bis auf wenige empirisch geleitete Systematisierungsversuche eher vage (u.a. Baumert et. al., o. J.; Fleischer, Koeppen, Kenk, Klieme & Leutner, 2013; Leuders & Sodian, 2013; Fleischer, Wirth, Rumann & Leutner, 2010; OECD, 2004; Billing D., 2003; Maag-Merki & Grob, 2001; Rychen & Salganik, 2001; Bennett, 2000; De la Harpe, Radloff & Wyber, 2000; Salganik, Rychen, Moser & Konstant, 1998; Clanchy & Ballard, 1995; Holmes, 1994). Dies ist insbesondere auch dem Umstand geschuldet, dass Definition und Operationalisierung je nach Untersuchungsinteresse, Forschungsfrage und empirischem Kontext variieren (Witmer, Storni & Schmid, 2008; Salganik, Rychen, Moser & Konstant, 1998). Selbst Mertens kam deshalb noch in den 80er Jahren zu dem Schluss, dass die Auseinandersetzung bisher „[e]igentlich nicht sehr viel mehr als eine abstrakte Diskussion, viel verbale Zustimmung mit wenig Umsetzung, überhaupt wenig Handfestes“ geleistet habe (Mertens 1988, S. 43).

Zum anderen sind als Konsequenz dieser Unklarheit auch Lehrmethoden bezüglich fächerübergreifender Problemlösekompetenzen bislang eher rar und schwierig umzusetzen (bspw. Lehre von Transfer). Daneben wird auch das Zusammenspiel zwischen kognitiven, metakognitiven und motivationalen *Skills* mit Problemlösen eher wenig in der Lehre berücksichtigt (Mayer, 1998). Ein Grund hierfür ist sicherlich, dass motivationale oder auch metakognitive *Skills* eher verborgen agieren und ihre Effekte nicht direkt beobachtbar sind (Schommer-Aikins, Duell & Hutter, 2005).

Diese Unschärfe in der Beschäftigung mit fächerübergreifendem Problemlösen stellt eine Herausforderung dar. Insbesondere auch deshalb, da Studienabsolventen nicht der Meinung sind, dass sie ausreichend kompetent sind, um mit fächerübergreifenden Problemen in ihrem künftigen

Berufsalltag umgehen zu können. So zeigte sich, dass Studierende verschiedener Studienfächer im nationalen und internationalen Kontext eine engere Verknüpfung ihres Studienfachs und der arbeitsbezogenen Praxis wünschen (Grützmaker, Ortenburger & Heine, 2011; Allen & van der Velden, 2007; Teichler & Schomburg, 2006; Borden & Rajackie, 2000). Es liegen weiterhin Befunde vor, dass sich viele Studierende nicht oder kaum unterstützt fühlen, was die Förderung von Problemlösefähigkeiten im Studium anbetrifft (bspw. Simeaner, Ramm & Kolbert-Ramm, 2010). Besonders interessant ist in diesem Zusammenhang auch, dass sich die Studierenden in den Magister- und Diplom-Studiengängen besser unterstützt fühlen als diejenigen Studierenden, die im Rahmen der Bachelor-Studiengänge besonders auf die künftigen beruflichen Anforderungen vorbereitet werden sollten (Multrus, 2011). Dabei ist die Thematik unzureichender Problemlösekompetenzen nicht unbedingt neu. Vielfach wurde bereits darauf hingewiesen, dass viele Studierende auf unterschiedlichen Ebenen ihres Studiums nicht in ausreichendem Maße in der Lage zu sein scheinen, kritisch zu denken und Probleme zu lösen (Sternberg, 1994). Als Konsequenz verwies Barnett (2000) bereits vor Jahren darauf, dass sich die Hochschulen an die veränderten Anforderungen anpassen müssen:

“Accordingly, a new responsibility is falling on universities to demonstrate that the education that they offer is likely to be adequate to the challenges [...]. It is a responsibility – and an educational project – that most universities and most curricula are failing to meet.” (Barnett, 2000, S. 265).

Die vorliegende Studie orientiert sich an dieser Forderung und greift die Desiderata im Zusammenhang mit fächerübergreifendem Problemlösen auf: Es wurde ein diagnostisches Instrument entwickelt, das auf aktuellen theoretischen und normativen Ideen fächerübergreifender Problemlösekompetenzen basiert und kognitive Ansätze zum Problemlösen berücksichtigt. Das Instrument will dementsprechend kognitive Kompetenzen des Problemlösens im fächerübergreifenden Kontext messen. Dabei wurden motivationale Faktoren ebenso wie metakognitive mit einbezogen, um diese Kompetenzen als behaviorale Kompetenzen im Sinne Weinerts (2001) zu verstehen und den Einfluss „verborgener“ personbezogener Faktoren auf die Problemlöseleistung mit zu berücksichtigen. Vor dem Hintergrund der bisherigen Annahmen und Befunde zur fächerübergreifenden Problemlösekompetenz der Studierenden fokussiert die Studie auch die Frage nach studierendenzentrierter Lehre (Star & Hammer, 2008; Biggs, 2003; Prosser & Trigwell, 1999) und thematisiert ebenso die Notwendigkeit der Vermittlung bestimmter Teilkompetenzen fächerübergreifenden Problemlösens.

Welche Fragestellung verfolgt die Studie?

In erster Linie fragt die Studie danach, wie fächerübergreifende Problemlösekompetenzen systematisch beschrieben und gemessen werden können. Im Weiteren liegt das Interesse in einer

gezielten Untersuchung der Ausprägung fächerübergreifender Problemlösekompetenzen bei Studierenden auf Bachelorniveau und der Frage von Unterschieden im Studienverlauf. Letztlich fragt die Studie danach, welche Bedingungen fächerübergreifender Problemlösekompetenz identifiziert werden können: Inwieweit nehmen motivationale, metakognitive und affektive Faktoren, die in der Person des Problemlösers liegen, Einfluss auf die Fähigkeit, fächerübergreifende Probleme lösen zu können? Inwieweit spielt die Konzeption des Problems eine Rolle? Und was ergibt sich daraus für die Lehre?

Die Arbeit umfasst fünf Teile: **Teil I** befasst sich mit der hochschulpolitischen Verortung des Konstrukts und stellt bisherige Ansätze zu seiner theoretischen und empirischen Fassbarkeit dar. Er stellt in einem kurzen Überblick die Diskussion um sogenannte *academic skills* dar und erläutert die Relevanz fächerübergreifender Problemlösekompetenzen in der Hochschulausbildung. Weiterhin wird die Debatte um den Kompetenzbegriff dargelegt. Der Kompetenzbegriff wird bislang uneinheitlich verwendet, weshalb es für die Studie unabdingbar ist, eine handhabbare Arbeitsdefinition vorzulegen, aus der ein überprüfbares Kompetenzkonzept abgeleitet werden kann. In **Teil II** wird die theoretische Verortung der Arbeit dargestellt. Die theoretische Fundierung des entwickelten Messinstrumentes entspannt sich zwischen den Konstrukten „Problemlösen“ und „fächerübergreifend“. Es werden Theorien und Befunde zum Problemlösen, zum Thema Transfer sowie zum Konzept des Fächerübergreifenden dargestellt. Ein Abriss über vorhandene diagnostische Instrumentarien zur Erfassung fächerübergreifender Problemlösekompetenzen bei Studierenden beschließt diesen Teil der Arbeit. **Teil III** stellt die Studie vor. Das entwickelte Instrument wird beschrieben und das der Studie zugrunde liegende methodische Design erläutert. Dabei wird auch die Überprüfung des Testinstrumentes im Rahmen der Pilotierung vorgestellt. Im Weiteren werden die Fragebögen beschrieben, die im Rahmen der Hauptstudie begleitend eingesetzt wurden. Die Ergebnisse der Studie werden in **Teil IV** dargestellt. Anschließend werden diese in **Teil V** vor dem Hintergrund der theoretischen Grundlage der Arbeit diskutiert und kritisch bewertet. Dabei werden auch mögliche Maßnahmen für die Lehre als Konsequenz aus den Befunden und Weiterentwicklungen des Messinstrumentes dargestellt. Die Arbeit schließt mit einer Zusammenfassung und einem Ausblick.

Teil I: Herausforderungen im Zusammenhang mit der Messung fächerübergreifender Problemlösekompetenz

1. *Academic skills* und fächerübergreifendes Problemlösen – Zur Ausgangslage der bildungspolitischen Diskussion und Einordnung des Untersuchungsgegenstandes

Die vorliegende Arbeit widmet sich einigen zentralen Fragestellungen der Diskussion um die sogenannte *employability* (Beschäftigungsbefähigung), bzw. *academic skills*, das heißt, an der Hochschule zu vermittelnden Kompetenzen:

- Im Hinblick auf eine theoretische Klärung fächerübergreifender Problemlösekompetenzen als Teilkomponente der Beschäftigungsbefähigung wird geklärt, wie fächerübergreifende Problemlösekompetenzen definiert werden können, welche Teilkomponenten eine Rolle spielen und mit welchen Methoden und mit welchem Instrument sie möglichst valide und reliabel gemessen werden können.
- In methodischer Hinsicht wird danach gefragt, wie fächerübergreifende Problemlösekompetenzen als Teilkomponente sogenannter generischer Kompetenzen zur Förderung von Beschäftigungsbefähigung systematisiert und modelliert werden können und wie diese Systematisierung in ein Messinstrument überführt und erprobt werden kann.
- In Bezug auf die Überprüfung von Kompetenzniveaus widmet sich die Arbeit der Frage, welche Leistungsunterschiede sich identifizieren lassen, wie sich fächerübergreifende Problemlösekompetenzen in Abhängigkeit von unterschiedlichen kognitiven Teilkompetenzen entwickeln, welche Bedeutung affektiv-motivationalen Facetten zukommt und wie zwischen erfolgreichen und weniger erfolgreichen Studierenden differenziert werden kann.

Wie lassen sich diese Fragen aus der Diskussion um Beschäftigungsbefähigung, bzw. *academic skills* herleiten? Warum ist eine Beschäftigung mit fächerübergreifenden Problemlösekompetenzen relevant und in welchen bildungspolitischen Zusammenhang muss die Beschäftigung mit fächerübergreifenden Problemlösekompetenzen eingeordnet werden?

Zur bildungspolitischen Einordnung der Thematik kann der Bogen weit zurück gespannt werden und der Bezug fächerübergreifender Problemlösekompetenzen als Teilkomponenten der sogenannten Schlüsselkompetenzen auf das Ende 1997 gestartete Projekt DeSeCo der OECD (*Definition and Selection of Competencies*) hergestellt werden, das in der Folge der Anfang 1997 etablierten PISA-Studien entwickelt wurde: Im Zusammenhang mit DeSeCo wurde danach gefragt, welche Schlüsselkompetenzen Bürgerinnen und Bürger grundsätzlich benötigen, um ein erfolgreiches Leben führen und eine gut funktionierende Gesellschaft aufrechterhalten zu können (Rychen & Salganik, 2003). Mit diesem Projekt sollte ein konzeptueller Referenzrahmen für die Kompetenzmessungen etabliert werden. Seit den ersten PISA-Studien in Deutschland im Jahr 2000 ist die Frage nach dem

Kompetenzerwerb in der gesellschaftlichen Debatte in Deutschland angekommen, seit dem Bologna-Prozess dringt die Frage in das Bewusstsein der deutschen Hochschulen.

Auf das Hochschulsystem bezogen berührt die Frage der Vermittlung fächerübergreifender Problemlösekompetenzen im Spezifischen sowie die Frage der Kompetenzanalyse im Allgemeinen im Kern ein bedeutendes, jedoch auch ein oft kontrovers diskutiertes Thema. Hinter der Frage, welche Kompetenzen Studierende haben sollten, steht letztlich die Frage danach, was genau ein Hochschulstudium ausmacht. Welche Kompetenzen sollen, bzw. können an einer Hochschule vermittelt werden? Wie können akademisch vermittelte Kompetenzen (*academic skills*) definiert werden? Im Folgenden wird dargestellt, welche Ansätze bislang zur Definition von *academic skills* vorliegen und inwiefern fächerübergreifendes Problemlösen darin einzuordnen ist. Einleitend wird in einem Überblick ebenso skizziert, an welchen Stellen im Zusammenhang mit der Debatte um *academic skills* Schwierigkeiten auftauchen und Kritik geäußert wird. Dies ist insofern von Bedeutung, da die Kompetenzdebatte durch unterschiedliche Zielsetzungen begründet wird und auf unterschiedlichen Ebenen auf Kritik stößt, die Berücksichtigung finden muss.

1.1 Fragestellungen im Zusammenhang mit der Diskussion um academic skills

In erster Linie zeigt ein Blick auf die Kompetenzdebatte, dass sie nicht nur deutschlandweit geführt wird und weiterhin auch keineswegs neu ist, wie mit ihrem Aufkommen im Zusammenhang mit dem Bologna-Prozess zu vermuten ist: Auf europäischer politischer Ebene besteht die Intention der vergleichenden Kompetenzanalyse mit Blick auf die gegenseitige Anerkennung von Qualifikationen bereits seit vielen Jahrzehnten und begann mit der Absichtserklärung der gegenseitigen Anerkennung von Befähigungsnachweisen in den Römischen Verträgen im Jahr 1957 (Bouder, 2006). In der Folge fand die Kompetenzdebatte 1998 im sogenannten Bologna-Prozess ihren gesetzmäßigen Auftakt, die Idee eines gemeinsamen europäischen Hochschulraumes wurde politisch umgesetzt, vergleichbare Bildungsabschlüsse wurden mit der Vergabe von Bachelor, Master und PhD eingeführt. In diesem Zusammenhang wurde die Definition von an der Hochschule zu vermittelnden Kompetenzen politisch festgeschrieben. Bereits 1988 wurde in der *Magna Charta Universitatum* auf Ebene der europäischen hochschulischen Bildungsinstitutionen in Bologna jedoch ebenso eine Erklärung auf Hochschulebene gesetzt. Dabei wurde insbesondere vor dem Hintergrund der politischen Diskussion der Grenzauflösung im Bildungssektor (*General Agreement on Trade in Services*, GATS) die Bedeutung der europäischen Universitäten und ihrer Autonomie definiert und die Mittel festgeschrieben, die zur Aufrechterhaltung ihrer Funktion als Vermittlerin von Wissen zur Bildung und Ausbildung zur Gewährleistung einer funktionierenden Gesellschaft notwendig sind.

Die Debatte um die Bedeutung (und Reform) von Hochschulen im Allgemeinen und den Kompetenzerwerb im Besonderen wird jedoch nicht nur europaweit, sondern auch weltweit geführt, die Implikationen des europäischen Bologna-Prozesses werden allseits diskutiert (bspw. Berdahl,

2010; Ilic, 2007; Terry, 2007). Wird der Rahmen zeit- und ortsbezogen schließlich noch weiter geöffnet, so zeigt sich, dass die Debatte um den Auftrag der Hochschulen im Grunde seit ihrer Etablierung als Bildungsinstitution immer wieder neu geführt wurde und im Zusammenhang mit ihrer politischen Legitimation und ihrer (zivil)gesellschaftlichen Bedeutung sicherlich immer wieder geführt werden wird (bspw. Ortega y Gasset, 1944; Jaspers, 1923).

Dass die Debatte um *academic skills* kontrovers diskutiert wird, ist jedoch nicht nur der Schwierigkeit der Thematik, sondern auch den Zielen geschuldet, die aktuell proklamiert werden. Welche sind dies und inwieweit führen sie zu Kritik?

Die aktuelle Frage nach Kompetenzen, die im Bereich der Hochschule erworben werden, soll erfolgen, um

- (1.) die Leistung im tertiären Bildungsbereich zu erhöhen,
- (2.) Kompetenzevaluierungen zu etablieren, um bildungspolitische Entscheidungen evidenzbasiert vornehmen zu können und
- (3.) mit internationaler Kompetenzforschung im Bereich der *Higher Education* mithalten zu können (Blömeke, Zlatkin-Troitschanskaia, Kuhn & Fege, 2013; auch Blömeke & Zlatkin-Troitschanskaia, 2013).

Am Beispiel des Projektes AHELO (*Assessment of Higher Education Learning Outcomes*, OECD), das Indikatoren des Kompetenzerwerbs an Hochschulen identifizieren will, werden die Ziele der Kompetenzmessung folgendermaßen formuliert: Auf Basis von Wirksamkeitsprüfungen sollen evidenzbasierte Aussagen zum *output* bzw. *outcome* einer heterogenen Hochschullandschaft gewonnen und für nachhaltige Entwicklungs- und Optimierungsmaßnahmen auf struktureller, organisatorischer und individueller Ebene genutzt werden (auch Blömeke, Zlatkin-Troitschanskaia, Kuhn & Fege, 2013; Blömeke & Zlatkin-Troitschanskaia, 2013; Braun, 2011). Anhand dieser Zielsetzungen wird deutlich, warum die Beschäftigung mit Kompetenzen im Hochschulbereich auf unterschiedlichen Ebenen auf Kritik stößt, die sowohl politischer als auch methodischer Natur ist und für deren Lösung bislang noch kein akzeptables Konzept vorliegt, geschweige denn ein Konsens, ob dieses Vorgehen nutzenbringend ist. Um einige Kritikpunkte zu nennen, die sich aus diesen Zielen ableiten lassen:

- Strukturell berührt die Frage nach der Kompetenzanalyse mit dem Ziel ihrer Festschreibung bspw. in Qualifikationsrahmen die bisherige Diversität der deutschen Hochschullandschaft. Sie stellt die Frage nach Unterschieden und Gemeinsamkeiten zwischen verschiedenen Bildungsformen, bspw. zwischen Berufsausbildung, dualem Hochschulstudium, Fachhochschulstudium und universitärem Studium. Insbesondere stellt sich die Frage, inwieweit nach der Einführung von Bachelor und Master mit den deutschlandweit bislang bildungspolitisch verankerten Unterschieden zwischen den Hochschultypen weiter umzugehen ist.

- Bildungspolitisch interessant ist die Frage nach der Definition von *academic skills* und ihrer Festsetzung in Qualifikationsrahmen, da Gestaltungsfreiheiten der Hochschulen, die einerseits im Zusammenhang mit Bologna (bspw. eigene Qualitätsentwicklung über die Systemakkreditierung) sowie mit Blick auf Konkurrenzfähigkeit und zunehmende (finanzielle) Alleinstellung der Hochschulen vergrößert werden sollen, mit Qualifikationsrahmen an anderer Seite ggf. wieder beschnitten werden könnten. Kritiker merkten zumindest in der Vergangenheit an, dass mit Qualifikationsrahmen die zu erwerbenden Kompetenzen ggf. zu spezifisch festgeschrieben und Gestaltungsfreiheiten wieder nivelliert werden könnten (bspw. DGB, 2009).
- Die wohl größte Kritik rührt aus dem Aspekt der Etablierung eines Hochschulmonitorings her, das heißt, eines „Hochschulpisas“. Was geschieht mit Hochschulen, die bei einem internationalen Kompetenzvergleich schlechter abschneiden? Relevanz besitzt dieser Punkt auch dahingehend, wenn bisherige Befunde eines „teaching to the test“, das heißt, eine Anpassung der Lehre an potenzielle Messinhalte, berücksichtigt werden (Nowakowski, Vervecken, Braun & Hannover, 2012): Ist eine Anpassung der Lehre an die Messinhalte eine (realistische) Gefahr?
- Neben diesem Auszug an politisch und gesellschaftlich motivierten Kritikpunkten stellt in methodischer Hinsicht eine valide und zuverlässige Modellierung und Erfassung akademisch vermittelter Kompetenzen sowie ihrer Bedingungen, Entwicklung und Wirkungen vor dem Hintergrund ihrer Multidimensionalität und -faktorialität hohe Ansprüche an die Forschungsmethodik, die bislang noch nicht eingelöst werden konnten.
- Die Problematik im Zusammenhang mit dem Konstrukt der *academic skills* ist dabei auch auf die Komplexität zurückzuführen, die akademisch erworbene Kompetenzen aufgrund der bisherigen Vielfalt an Studienmodellen, Ausbildungsstrukturen und Lehrangeboten auszeichnet. Welches Kriterium soll deshalb zur Messung herangezogen werden, bzw. inwieweit kann das gemessene Ergebnis mit der Lehre in Verbindung gebracht werden, wenn die Heterogenität kaum kontrolliert werden kann? Sollen die vermittelten Fachinhalte, zukünftige Berufsanforderungen oder aus Qualifikationsrahmen abgeleitete Kompetenzen zur Ableitung von Kriterien dienen?
- Des Weiteren stellt sich die Frage, inwieweit aus teilweise sehr breiten beruflichen Einsatzbereichen und Anforderungen im Studium zu vermittelnde Kernkompetenzen operationalisiert werden können. Weiterhin sind berufliche Kompetenzen in vielen Domänen (bspw. Soziologie, Philosophie) nur schwer erfassbar bzw. definierbar und unterliegen ebenso einer ständigen Weiterentwicklung.

Um diese Kritikpunkte zusammenzufassen: Auf der einen Seite stehen Kritiker, die auf politische Gefahren und methodische Schwierigkeiten verweisen. Auf der anderen Seite stehen die Befürworter,

die hochschulische Kompetenzen messbar und über Ländergrenzen hinweg vergleichbar machen wollen. Auch wenn zahlreiche bedeutsame Fragen noch nicht beantwortet sind – warum ist eine Auseinandersetzung mit der Thematik der Kompetenzanalyse dennoch so relevant?

Der Ansatzpunkt der vorliegenden Arbeit, der im Zusammenhang mit der Kompetenzanalyse kaum genannt wird, orientiert sich an den Lehrenden und Studierenden:

So zeigt sich einerseits, dass viele Studienabsolventen der Meinung sind, dass sie nicht ausreichend kompetent sind, um mit Problemen in ihrem künftigen Berufsalltag umgehen zu können (Grützmaker, Ortenburger & Heine, 2011; Multrus, 2011; Simeaner, Ramm & Kolbert-Ramm, 2010; Allen & van der Velden, 2007; Teichler & Schomburg, 2006; Borden & Rajeckie, 2000). Auf der anderen Seite verweisen auch Lehrende seit Jahren darauf, dass Studierende unzureichend in der Lage sind, sich mit Problemen auseinanderzusetzen, die Hochschule sich deshalb als Bildungsinstitution verändern müsse (Barnett, 2000; Sternberg, 1994). Die aktuelle Kompetenzdebatte setzt sich im Kern mit diesen Fragen auseinander, wenn sie auch von anderen Zielen ausgeht: Welche Kompetenzen müssen an der Hochschule gelehrt und gelernt werden? Daran schließt sich die Frage an, welche Kompetenzen gemessen werden müssen, bzw. gemessen werden können, um nach der Generierung von Daten und Kennwerten zum Kompetenzerwerb die erworbenen Kompetenzen einordnen, Lehre verbessern, gute Lehrergebnisse aus einem Bereich in einen anderen transferieren zu können, oder auch die Faktoren guter Lehre ausmachen zu können.

1.2 Ansatzpunkte zur Operationalisierung von academic skills und Einordnung fächerübergreifender Problemlösekompetenzen

In erster Linie liegen bislang eher deskriptiv orientierte Beschreibungen von *academic skills* vor. So sollen Hochschulen „reflective practitioners, social critics and good citizens“ ausbilden (Star & Hammer, 2008, S. 2), im Vergleich zur Berufsausbildung soll die Hochschulausbildung sich auszeichnen als

“[a]cademic qualification [...], theoretical approach to systematic problem-solving that is not contained in the degrees and training programmes of vocational education. [...] Competent university graduates help companies by challenging old patterns of thought and habit and thus adjust them for the demands of a dynamic, international labour market” (Zervakis, 2008, Online Ressource).

Weiter wird definiert, dass „Absolventen befähigt [sein sollen; Anm. d. Autorin], eine qualifizierte Beschäftigung aufzunehmen, zu halten und sich neue Beschäftigungsfelder zu erschließen“ (BDA, BDI & HRK, 2008, S. 5). Diese Forderung wird unter dem Begriff der Ausbildung an der Hochschule hin zur *employability*, das heißt, zur Beschäftigungsbefähigung, zusammengefasst. Starr und Hammer (2008) verweisen in diesem Zusammenhang auf die Notwendigkeit einer Verschiebung von Inhalten

hin zu Prozessen. Das heißt, zur Operationalisierung von Kompetenzen sollte die Frage leitend sein, was Absolventen mit ihrem Wissen anfangen können. Sie schließen deshalb, dass Absolventen befähigt werden müssen “to understand, assign significance and interpret through developing intellectual skills like problem-solving, logical thinking and information gathering, with a subordination of the acquisition of knowledge and the facts that underpin the discipline.” (Barrows, nach Starr & Hammer, 2008, S. 8)

Im Weiteren werden *academic skills* definiert, indem zwischen generischen und fachspezifischen Kompetenzen unterschieden wird. Generische Kompetenzen bezeichnen dabei fächerübergreifende Kompetenzen, die generell für ein akademisches Abschlussniveau gelten. Sie umfassen instrumentelle, interpersonelle und systemische Kompetenzen wie bspw. Analyse- und Synthesefähigkeiten, die Fähigkeit, eigenes Handeln reflektieren, mit komplexen und neuartigen Problemstellungen umgehen, sowie wissenschaftliche Methoden systematisch einsetzen zu können. Fachspezifische Kompetenzen beziehen sich auf Qualifikationsanforderungen für eine fachliche Disziplin und umfassen Wissen, Verstehen und Anwenden sowie Wissensvertiefung und Wissensverbreiterung (Qualifikationsrahmen für deutsche Hochschulabschlüsse, 2005). Sie sind unerlässlich für eine Fachrichtung.

Zahlreiche nationale und internationale Befragungen unterstreichen die Relevanz dieser Teilkompetenzen bzw. Kompetenzbereiche, auch wenn Begrifflichkeiten unterschiedlich oder nicht klar voneinander getrennt verwendet werden: So wird bspw. von Berufsverbänden zur Vermittlung von Beschäftigungsbefähigung gefordert, fachwissenschaftliche und methodische Kompetenzen, soziale Kompetenzen sowie personale Kompetenzen zu lehren (BDA, BDI & HRK, 2008). Neben einer eher lexikalischen Auflistung von Kompetenzen, die zur Beschäftigungsbefähigung relevant sein könnten erfolgte auch der Versuch einer systematischen Analyse über unterschiedliche methodische Herangehensweisen wie bspw. die Analyse von Stelleninseraten (Dietzen, 1999; Spencer & Spencer, 1993), nationale und internationale Unternehmensbefragungen (Deutscher Industrie und Handelskammertag, 2004; Bennett et. al., 2000; Manninen & Hobrough, 2000; Dietzen, 1999; Austin et.al., 1997; Stasz, 1996; Tate & Thompson, 1994; Borthwick, 1993; Everwijn et. al., 1993; Spencer & Spencer, 1993; France, Vincens & Chirache, 1992; Green 1990) sowie faktorenanalytische Clusterungen gefundener Kompetenzen (Klemp & McClelland, 1986). Insgesamt überwiegen jedoch bislang deskriptive Beschreibungen von zu erwerbenden *academic skills*. Dabei zeigt sich, dass es zwar gewisse Variationen je nach Kontext des jeweiligen Berufs gibt, insgesamt aber den Kompetenzbereichen „Kommunikationsfähigkeit“, „Teamwork“ und „Problemlösen“ die höchste Wertigkeit am Arbeitsplatz zugeschrieben wird (Bennett et. al., 2000; Stasz et. al., 1996).

Wie wurden diese deskriptiven Ansätze weiter versucht zu validieren und operationalisieren?

Ein bedeutsamer Ansatzpunkt ist die Berücksichtigung der Hochschulabsolventen selbst. Über ihre subjektiven Vorstellungen können Hinweise darauf abgeleitet werden, inwieweit die Hochschulausbildung diejenigen Kompetenzen vermittelt, die ihrer Ansicht nach zur erfolgreichen

Berufsausübung als hilfreich angesehen werden. Zahlreiche nationale sowie internationale Befragungen geben Auskunft darüber, inwieweit Absolventen mit ihren erworbenen Kompetenzen zufrieden sind (bspw. *Kooperationsprojekt Absolventenstudien*, KOAB, seit 2007; *Careers after Higher Education: a European Research Study*, CHEERS I & II, 1997-2006; *Arbeitsituation von Universitäts- und FachhochschulschulabsolventInnen österreichischer Hochschulen*, ARUFA, 2009-2011; für einen Überblick Reinfeldt & Frings, 2003; auch Guggenberger et al., 2007; Schaeper & Briedis, 2004). Eine Vielzahl an hochschuleigenen Absolventenbefragungen ergänzen diese Studien. Zur Frage besonders bedeutsamer Kompetenzen, die im Hochschulstudium vermittelt werden müssen, wurden bspw. im Rahmen des EU-Projektes REFLEX (*Hochschule, Qualifikation und Berufstätigkeit in Europa*, 2004-2008) zirka 36'000 europäische Hochschulabsolventen fünf bis sechs Jahre nach Studienabschluss zu ihren Berufserfahrungen befragt. Ergebnis war, dass für den deutschsprachigen Raum den Kompetenzfeldern „Organisationsfähigkeit“, „Problemlösekompetenz“, „Kommunikations- und Sozialkompetenz“ in dieser Reihenfolge die größte Bedeutung für den Arbeitsalltag zugesprochen wurde (Strub, Dubach & Oesch, 2008). Trotz der Vielzahl an Befragungen, der Heterogenität der Vorgehensweisen und Unterschiedlichkeit der Befragten wurden damit zwischen Arbeitgebern und künftigen Arbeitnehmern übereinstimmende Kompetenzfelder gefunden.

Die Bedeutsamkeit der Beschäftigung mit fächerübergreifendem Problemlösen basiert dementsprechend darauf, dass die Kompetenz, in neuen und unbekanntem Situationen Probleme lösen zu können als generische Kompetenz eine für die spätere Berufspraxis sowohl von Absolventen als auch Arbeitgebern genannte bedeutsame Fähigkeit darstellt. Besonders bedeutsam wird Problemlösen im Allgemeinen und fächerübergreifendes Problemlösen im Besonderen jedoch auch deshalb, da sie als zentrale Fähigkeit begriffen wird, neues Wissen zu erschließen (Baumert, Klieme, Neubrand, Prenzel, Schiefele, Schneider, Stanat, Tillman, Weiß, 2001).

Für die weitere Operationalisierung der generischen und fachspezifischen Kompetenzen im Allgemeinen sowie der fächerübergreifenden Problemlösekompetenz im Spezifischen fehlen jedoch bislang valide Kompetenzmodelle, das heißt, strukturierte und auf das Hochschulniveau angepasste Modelle die sich auf die jeweilige Domäne beziehen (u.a. Blömeke, Zlatkin-Troitschanskaia, Kuhn & Fege, 2013; Fleischer, Koeppen, Kenk, Klieme & Leutner, 2013; Leuders & Sodian, 2013). Erste Ansätze gibt es bspw. in Form des „Qualifikationsrahmens für Deutsche Hochschulabschlüsse“. Darin wurden zu erwerbende Kompetenzen („Wissen und Verstehen“ und „Können“) und formale Aspekte (z. B. Zugangsvoraussetzungen, Dauer) gestuft nach Abschlüssen (BA, MA, Doktorat) formuliert (HRK, KMK & BMBF, 2005). Sie stellen einen ersten Ansatz vergleichender Kompetenzmessung dar, sind jedoch nicht empirisch fundiert und besitzen, bis auf den Ansatz der Sozialen Arbeit (Bartosch, Maile & Speth, 2008), keinen Domänenbezug. Evidenzbasierte Kompetenzmodelle im Hochschulbereich sind bislang die Ausnahme (bspw. Baumert & Kunter, 2006). Das bedeutet, dass ein

robustes Forschungsdesign zur Erfassung von Kompetenzniveaus entwickelt werden muss, um Indikatoren guter Lehre auszumachen und Lehre verbessern zu können. Dabei sollte sich die Konstruktion von Kompetenzmodellen auf empirisch gestützte Theorien zu kognitiven Prozessen beziehen, wobei die Validität des Modells und seiner Operationalisierungen durch eine Orientierung sowohl an geeigneten kognitionspsychologischen als auch an auf die Lehre bezogenen Analysen gesichert werden sollten (Leuders & Sodian, 2013).

Diesem Vorhaben widmet sich mit den eingangs genannten Forschungsfragen die vorliegende Arbeit. Um im Folgenden zu einer geeigneten Arbeitsdefinition von fächerübergreifender Problemlösekompetenz zu gelangen wird im Folgenden die Diskussion um den Kompetenzbegriff skizziert, wobei nach einer Darstellung bislang vorliegender Systematisierungsansätze die für die Arbeit rahmende kognitive Kompetenzdefinition begründet wird.

2. Der Kompetenzbegriff: Konjunktur und Problematik

Die in erster Linie domänenspezifisch ausgerichtete Kompetenzdebatte dominieren sowohl im Anwendungs- wie auch im Forschungsbereich Fragen nach der Modellierung von Kompetenzen, ihrer Messung sowie ihrer Förderung. Dies sind Fragestellungen, zur deren Beantwortung jedoch zuerst definiert werden muss, was es bedeutet, kompetent zu sein. Dies ist umso wichtiger, da aufgrund der inflationären Begriffsverwendung und der relativ willkürlichen Verwendung und Operationalisierung des Kompetenzkonstrukts, wie sie auch der bildungspolitische Abriss zeigte, die Gefahr besteht, sozial definierte Konstrukte in die wissenschaftliche Diskussion zu übernehmen (Klieme et. al., 2001a; auch Rosenstiel & Erpenbeck, 2007; Klieme, Funke, Leutner, Reimann, Wirth & Weinert, 2001; Barnett, 1994).

Was bedeutet kompetentes Verhalten? Um Kompetenz nachzuweisen, genügt es im Regelfall nicht, deren Vorhandensein zu behaupten (Gnahn, 2010). Kompetentes Verhalten zeigt sich vielmehr in der Bewältigung bestimmter Situationen. Kompetenz ist dementsprechend erst aus der Realisation der Disposition erschließbar und evaluierbar (Erpenbeck & Rosenstiel, 2007; Klieme & Hartig, 2007).

Um das Kompetenzkonstrukt näher zu erläutern und die Problematik im Umgang mit dem Begriff zu skizzieren, werden im Folgenden die wesentlichen Aspekte der Diskussion hierzu vorgestellt. Einen essentiellen Beitrag legten Weinert (2001), Klieme und Hartig (2007) sowie Hartig und Klieme (2006, 2007) und Koeppen, Hartig, Klieme und Leutner (2008) vor. Sie haben die wissenschaftliche Diskussion um den Kompetenzbegriff systematisch historisch nachgezeichnet, die unterschiedlichen Schwerpunkte der Begriffshistorie mit Blick auf die Sozial- und Erziehungswissenschaften herausgearbeitet und Konsequenzen, Möglichkeiten sowie Anforderungen an eine Kompetenzdiagnostik dargelegt. Ihre Herangehensweise setzt sich von der Vielzahl an Publikationen

zum Kompetenzbegriff nicht nur aufgrund ihrer vergleichenden Analyse ab, sondern erlaubt durch die Erarbeitung einer als Arbeitsgrundlage geeigneten Definition des Kompetenzbegriffs und den Hinweis auf Anforderungen an eine Kompetenzanalyse wertvolle Ansatzpunkte. Im Folgenden werden bei der Darstellung von Ansätzen zur Systematisierung von Kompetenz insbesondere auch ihre Hauptargumente nachgezeichnet. Der Schwerpunkt wird darauf gelegt, die der Arbeit zugrundeliegende funktional-pragmatische Kompetenzdefinition zu begründen und die daraus resultierenden Anforderungen an eine Kompetenzdiagnostik zu erläutern.

2.1 Systematisierungsansätze

Ausgehend vom Alltagsverständnis werden unter dem Begriff sowohl Fähigkeiten, als auch Zuständigkeiten und Bereitschaften subsummiert. Als Synonyme werden Sachverstand, Begabungen, Qualifikationen sowie Entscheidungsbefugnisse genannt (Duden, 2013). Diese Allgegenwärtigkeit und Vielschichtigkeit des Begriffs führt dazu, dass auch in der wissenschaftlichen Beschäftigung unter dem Kompetenzkonstrukt Unterschiedliches verstanden wird. Die in der Diskussion oftmals vorherrschende Unschärfe und Simplifizierung des Kompetenzbegriffs (Arnold, 2002) ist allerdings sicherlich auch darauf zurückzuführen, dass der aktuelle Diskurs nicht wissenschaftsintern angestoßen wurde, sondern vielmehr aus gesellschaftlichen Entwicklungen entstand (Geißler & Orthey, 2002). Die gegenwärtige Diskussion um den Begriff scheint zudem weniger ein Systematisierungs-, sondern vielmehr ein praktisches Lösungsinteresse zu beherrschen: Wer im Diskurs nicht abgehängt werden wolle, dürfe beim Trend zur Kompetenz nicht abseits stehen, so spitzen es Geißler und Orthey zu (2002; auch Koeppen, Hartig, Klieme, & Leutner, 2008; Arnold, 2002; Brödel, 2002; Orthey 2002).

Im Zuge der wissenschaftlichen Beschäftigung mit Kompetenz nahm die Popularität des Begriffs vor allem mit Chomskys sprachwissenschaftlich verortetem Kompetenz-Performanz-Modell ihren Anfang, in der Folge wurde der Kompetenzbegriff vor allem in den Sozial- und Erziehungswissenschaften immer wieder neu gefasst (Ashworth & Saxton, 1990; auch Klieme & Hartig, 2007; Geißler & Orthey, 2002; Orthey 2002; Arnold, 2002; Brödel, 2002). Unterschiedliche Konjunkturen, Traditionen und Herangehensweisen nahmen Einfluss und führten zu einer zunehmenden Komplexität des Begriffs sowie zu oftmals widersprüchlichen Kompetenzkonzepten (Ashworth, Saxton, 1990; für einen Überblick über verschiedene Konzepte Weinert, 2001):

„[...] [I]n all of these disciplines [developmental sciences, psychology, linguistics, sociology, political science, economics], competence is interpreted as a roughly specialized system of individual and/or collective abilities, proficiencies, or skills, that are necessary or sufficient to reach a specific goal.“
(Weinert, 2001, S. 5)

Es entwickelten sich zwar gemeinsame Vorgehensweisen zur Systematisierung des Begriffs, beispielsweise über Bedarfsanalysen (bspw. Einleitung dieser Arbeit) oder über geschichtliche (bspw. Vonken, 2005; Huber, 2001) und begriffskulturelle Einordnungen (bspw. Erpenbeck & Rosenstiel, 2007). Einteilungsversuche resultierten in Systematiken, die zum Beispiel Kompetenztypen, Kompetenzklassen und Kompetenzgruppen unterscheiden (bspw. Erpenbeck & Rosenstiel, 2007). Eine Fassung des Kompetenzbegriffs wurde dabei jedoch bis heute mit unterschiedlichen theoretischen, normativen und empirischen Bedeutungen versehen und führte keineswegs zu einer allgemein verbindlichen Definition (für eine Übersicht Weinert, 2001). Darauf verweisen auch Klieme und Hartig (2007), wenn sie die Komplexität des Begriffs versuchen abzubilden:

„Kompetenz [...] im situativen Bewältigen von Anforderungen (in der ‚Performanz‘ des Handelns), wird aber als Disposition interpretiert. [...] Kompetenz [ist] kontextualisiert und spezifisch, aber auf Transfer und Verallgemeinerung angelegt. Kompetenz bezieht sich auf Handlungsvollzüge [...] [auf] die ihnen zugrundeliegenden mentalen Prozesse und Kapazitäten, zu denen Kognition, Motivation und Volition bzw. Wissen und Können gehören. Im wissenschaftlichen Kontext verbindet sich mit dem Kompetenzkonstrukt die Erwartung, dieses Wissen und Können in ökologisch valider Weise beschreiben und messen zu können. Im Bildungskontext verbindet sich ‚Kompetenzorientierung‘ zunächst ganz allgemein mit dem geschärften Blick auf die tatsächlich erreichten Lernergebnisse. Kompetenzorientierung bedeutet hier auch Wissen und Können so zu vermitteln, dass keine ‚trägen‘ und isolierten Kenntnisse und Fähigkeiten entstehen, sondern anwendungsfähiges Wissen und ganzheitliches Können, das z.B. reflektive und selbstregulative Prozesse einschließt.“ (Klieme & Hartig, 2007, S. 13)

Die Problematik der Systematisierung von Kompetenz basiert dementsprechend auf seiner Komplexität und scheint sich zudem auch als Theorie-Praxis-Problem darüber zu entspannen, dass sich Konzepte der beruflichen Handlungskompetenz und Konzepte der funktionalen und fachbezogenen Kompetenz gegenüberstehen. Die Beschäftigung mit dem Kompetenzbegriff führte demzufolge bis heute lediglich dazu, dass allenfalls eine abstrakte Leitvorstellung zum Begriff besteht (Klieme & Hartig, 2007): Der Begriff Kompetenz verweist auf Qualitäten des menschlichen Tuns und Denkens und ist dementsprechend nur schwer verzichtbar, wenn es darum geht, menschliches Handeln zu fassen.

Heißt dies, dass wir *Occam's razor* ansetzen müssen und eine Aufgabe des Kompetenzbegriffs sinnvoll ist, vor dem Hintergrund einer scheinbar unvereinbaren Definitionsvielfalt aufgrund der Komplexität des Konstrukts? Dies kann keine Lösung sein. Der Kompetenzbegriff bietet hilfreiche Ansatzpunkte und verbindende Elemente (Erpenbeck & Rosenstiel, 2007; Sarges, 2001). Zentral für

den Kompetenzbegriff ist vielmehr, dass er die Betrachtung von Fähigkeiten erlaubt, die mit Blick auf Anforderungen relevant werden, wobei weiterhin unterstellt wird, dass sie in einer Vielzahl von Situationen anwendbar sind. Das heißt, wer kompetent ist, verfügt nicht nur über träges, sondern anwendbares Wissen. Der Begriff der Kompetenz steht dementsprechend dafür, dass Handlungsfähigkeit als Kriterium gilt (Brödel, 2002).

Im Folgenden wird der Ansatz für die Definition von Kompetenz vorgestellt, der als Grundlage für diese Arbeit herangezogen wurde. Dabei wird weniger einem konditionalen Ansatz gefolgt und die Frage gestellt, welche Kompetenzen grundsätzlich identifiziert und definiert werden können. Um zu einer Operationalisierung fächerübergreifender Problemlösekompetenzen zu gelangen, wird einem funktional-pragmatischen Ansatz gefolgt (Weinert, 2001).

2.2 Das funktionale Kompetenzkonzept der Psychologie

In den Sozialwissenschaften war es Chomsky, der die Popularität des Begriffs und die weitere wissenschaftliche Beschäftigung mit dem Kompetenzkonstrukt etablierte, in der Folge entwickelten sich innerhalb der Sozialwissenschaften unterschiedliche Herangehensweisen zur Definition von Kompetenz (Klieme & Hartig, 2007; Weinert, 2001). Chomskys technischer Kompetenzbegriff zur Beschreibung des allen Menschen zugrundeliegenden kognitiven Systems des sprachlichen Handelns ist als abstraktes Modell jedoch ebenso wenig wie die in dieser Tradition stehenden Kompetenzmodelle (beispielsweise die computerbasierten Simulationsmodelle der Kognitionswissenschaften) darauf ausgelegt, die Auseinandersetzung des Individuums mit seiner Umwelt zu untersuchen. Vielmehr soll ein ideales System von Kompetenzen beschrieben werden, das zwar Verhalten, beziehungsweise Performanz erzeugt, damit aber nicht identisch ist (Kurtz & Pfadenhauer, 2010). Im Rahmen der Psychologie kritisierten White (1959) und später auch McClelland (1973) diese Utopiedefinition: Für beide war ausschlaggebend, dass die Kompetenzkonzepte keine Erklärung lieferten, wie sich eine wirkungsvolle Interaktion des Individuums mit der Umwelt entwickelt. Ihre Forderung war, der Kontextabhängigkeit menschlichen Handelns vermehrt Rechnung zu tragen. McClellands Forderung bezog allerdings auch mit ein, dass anhand kompetenzorientierter Testung eine bessere Passung zwischen Testanforderungen und realen, beziehungsweise beruflichen Situationen erreicht werden kann: „Do intelligence tests tap abilities that are responsible for job success?“ (1959, S. 3) Weiterhin sollten auch individuelle Leistungsunterschiede besser erklärt und vorhergesagt werden können. Zwar präziserte McClelland (ebd.) diese jobrelevanten Kompetenzen nicht weiter, mit White und McClelland erhielt die Kompetenzdebatte jedoch ihren Bezug zu realen Anforderungen: Das heißt, nicht nur Wissen und Fähigkeiten werden relevant, sondern auch deren erfolgreiche Anwendung in unterschiedlichen und neuen Situationen. Damit etablierte sich die Leitlinie für die weitere Kompetenzdebatte und die

Kontextspezifität von Kompetenzen hielt Einzug. Kompetenzen werden demzufolge bis heute als „realized abilities“ verstanden (Connell, Sheridan & Gardner, 2003, S. 142) und definiert als “[...] context-specific [...] dispositions that are acquired and needed to successfully cope with certain situations or tasks in specific domains” (Koeppen, Hartig, Klieme, & Leutner, 2008, S. 62). Das bedeutet, im pädagogisch-psychologischen Bereich sind Kompetenzen meist auf spezifische Kontexte bezogen und berücksichtigen demzufolge sowohl personen- als auch situations- oder domänenspezifische Faktoren (Hartig & Klieme, 2006; Weinert, 2001).

Allerdings bezieht die aktuelle Definition mit ein, dass die Kontextspezifität von Kompetenzen zwischen fachspezifischen und fachübergreifenden Kompetenzen variieren kann (Transferbreite, Weinert, 2001). Der Fokus auf die Kontextspezifität verweist auch auf einen weiteren für den Kompetenzbegriff leitenden Aspekt. Er verdeutlicht, dass Kompetenzen erlernbar, das heißt, durch die Umwelt beeinflussbar sind. Durch Training, Intervention und Praxis kann Kompetenz bis zur Expertise weiterentwickelt werden (Mayer, 2003).

Der Rückgriff auf die Geschichte des Kompetenzbegriffs zeigt, welche Leitlinien zur Kompetenzdefinition zugrunde zu legen sind: Kompetenzen werden als erlernbare kontextspezifische Leistungsdispositionen verstanden, die sich auf Anforderungen in bestimmten realen Situationen beziehen. Dabei ist die Breite der Domänen und Situationen variabel (auch Klieme & Hartig, 2007).

3. Kompetenz als Bewältigung von Handlungsanforderungen und Konsequenzen für die Messung des Kompetenzkonstrukts

Wird eine präzise Definition angestrebt, so ist ein wichtiger Schritt die Abgrenzung zu verwandten Konzepten. Dementsprechend müssen vor allem die Begriffe Qualifikation, Wissen und Intelligenz sowie Fähigkeiten und Fertigkeiten vom Kompetenzbegriff differenziert betrachtet werden (Frey, 2006; Vonken, 2005; Orthey & Geißler, 2002; Klieme et. al., 2001; Weinert, 2001). Für eine Kompetenzanalyse bleibt im Weiteren ebenso die Frage zu beantworten, wie Kompetenzen präzise erfasst werden können.

3.1 Eine pragmatische Arbeitsdefinition

So wird Kompetenz darüber definiert, dass sie stärker als Qualifikation sowohl an das Subjekt als auch den jeweiligen Kontext gekoppelt ist. Das heißt, Kompetenzen werden auf Basis eines Impulses von außen, einer Anforderung und der Einbeziehung des eigenen Handlungs- und Verhaltensrepertoires in Handlungssituationen aktualisiert. Im Gegensatz zur allgemeinen Intelligenz ist Kompetenz erlernbar und kontextabhängig. Intelligenz als zeitlich stabiles Konstrukt steht Kompetenz auch insofern

gegenüber, dass Kompetenz durch Erfahrung im Umgang mit Situationen und Anforderungen erworben und zur Expertise weiterentwickelt werden kann. Mit Blick auf die Binnenstruktur ist Kompetenz abhängig von den jeweiligen Anforderungen und Situationen, die bewältigt werden sollen, das heißt, sie ist bereichsspezifisch. Die Struktur der allgemeinen Intelligenz hingegen ergibt sich aus bereichsübergreifend relevanten kognitiven Fähigkeiten (für eine Diskussion Hartig & Klieme, 2006). Letztlich impliziert Kompetenz im Gegensatz zum Wissensbegriff Wissenserwerbs- und Wissensanwendungsaspekte und verbindet dementsprechend Handeln und Können. Diese letztgenannte Perspektive von Kompetenz als Fähigkeiten zur Bewältigung von Handlungsanforderungen ist es jedoch, die nicht nur eine Definition von Kompetenzen sondern auch ihre empirische Überprüfung erschwert, da kognitive, motivationale, volitionale und soziale Aspekte sowie ihre Beziehung zueinander im Rahmen des Kompetenzbegriff abgebildet werden. Weinert (2001) löst mit einer pragmatischen Definition diese Problematik allerdings auf. Zwar ist auch er anfangs uneindeutig, wenn er formuliert:

„Kompetenzen sind die bei Individuen verfügbaren oder erlernbaren kognitiven Fähigkeiten und Fertigkeiten um Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (S. 27f.)

Er schlägt im Weiteren jedoch vor, das Kompetenzkonstrukt im Rahmen empirischer Untersuchungen als rein kognitives Konstrukt zu untersuchen. Dementsprechend sollten kognitive Kompetenzkomponenten getrennt von motivationalen, volitionalen und sozialen erfasst werden. Erst diese Definition und Vorgehensweise ermöglicht eine präzise Operationalisierung und erlaubt, die Wechselwirkungen zwischen den Komponenten empirisch nachzuzeichnen (Weinert, 2001). Auch theoretisch legitimiert Weinert diese Trennung und bietet anhand der Trennung von Kompetenz und Handlungskompetenz eine Lösung über eine Reduktion des Kompetenzkonstrukts an. Die im Rahmen der Handlungskompetenz mit einbezogenen motivationalen und volitionalen Dispositionen sind zwar als Management- bzw. Regulationskomponenten im situativen Kompetenzvollzug nicht von kognitiven Kompetenzkomponenten zu trennen. Kompetenz ist allerdings als kognitives Konstrukt funktional-pragmatisch zu definieren und von Handlungskompetenz zu trennen. Dementsprechend definieren Klieme et. al. (2001d) Kompetenz letztlich als

„[...] funktional bestimmte, auf bestimmte Klassen von Situationen und Anforderungen bezogene kognitive Leistungsdispositionen, die sich psychologisch als Kenntnisse, Fertigkeiten, Strategien, Routinen oder auch bereichsspezifische Fähigkeiten beschreiben lassen.“

Kompetenzen sind demzufolge“ [...] Systeme aus spezifischen, prinzipiell erlernbaren Fertigkeiten, Kenntnissen und metakognitivem Wissen, die es erlauben, eine Klasse von Anforderungen in bestimmten Alltags-, Schul- oder Arbeitsumgebungen zu bewältigen.“ (Klieme et al., 2001d, S.182).

Das Verhältnis von kognitiver Kompetenz, Fähigkeiten und Fertigkeiten lässt sich demzufolge darüber beschreiben, dass kognitive Kompetenzen als Bündel von Fähigkeiten und Fertigkeiten angesehen werden (Weinert, 2001; Klieme et. al. 2001d). Fähigkeiten gelten dabei für die hierarchisch übergeordneten Kompetenzen als Indikatoren und können ihrerseits ausschließlich über Fertigkeiten gemessen, diagnostiziert und letztlich praktisch definiert werden (Frey, 2006). Kompetenz ist damit die Summe der Fähigkeiten und Fertigkeiten sowie der motivationalen, volitionalen und sozialen Bereitschaften eines Individuums, wie sie zur Lösung von Problemen in variablen Situationen benötigt werden. Das heisst, eine Kompetenz kann auch in Teilkompetenzen im Sinne derjenigen Fähigkeiten und Fertigkeiten sowie Bereitschaften gegliedert werden, wie sie zur Lösung von Problemen in bestimmten Teilbereichen benötigt werden. Die Grenze ihrer Feingliedrigkeit ist dann erreicht, wenn zur Lösung eines Problems nur noch einzelne Fähigkeiten oder Fertigkeiten oder Bereitschaften benötigt werden. Eine einzelne Fähigkeit oder Fertigkeit stellt keine Teilkompetenz mehr dar, wie Neumann (2013) betont. Ihre Begrenzung ist durch die Forderung nach Erlernbarkeit gegeben.

3.2 Anforderung an die Kompetenzmessung

Diese kognitive Kompetenzdefinition ermöglicht eine klare Orientierung zur Operationalisierung. Wie können Kompetenzen jedoch präzise gemessen werden? Problematisch ist eine Messung vor allem, da es eine Herausforderungen ist, für eine präzise Kompetenzmodellierung geeignete psychometrische sowie statistische Modelle zu entwickeln (Klieme & Hartig, 2007; auch Fleischer et. al., 2013). Ein detaillierter Abriss über die zur Messung von Kompetenzen geeigneten bzw. wünschenswerten statistischen Modelle würde an dieser Stelle zu weit führen. Hierfür kann auf Embretson (2006) sowie Hartig und Klieme (2006) oder auch Klieme und Hartig (2007) verwiesen werden. Allerdings muss mit Blick auf das dieser Arbeit zugrundeliegende Testinstrument zur Messung fächerübergreifender Problemlösekompetenzen die Frage geklärt werden, an welchen Ansprüchen sich Kompetenzdiagnostik messen lassen muss. Zudem soll dargestellt werden, welche Messmodelle vorliegen, die als Orientierung für die Entwicklung eines psychometrischen Modells der Kompetenzmessung dienen können.

Aus der Definition von Kompetenzen ergibt sich für ihre Messung, dass sowohl Kompetenz als auch Handlungskompetenz grundsätzlich Zuschreibungen aufgrund der Beobachtung von als kompetent definiertem Verhalten sind. Erst die konsistente Zusammenfassung von Beobachtungen über unterschiedliche Aufgaben hinweg, beziehungsweise in verschiedenen Situationen, erlaubt eine individuelle Ausprägung von Kompetenzen empirisch möglichst eindeutig zu bestimmen. Ein

entsprechendes theoretisches Kompetenzmodell ist demzufolge ein Modell darüber, wie unterschiedliche kognitive Kompetenzen in bestimmten situativen Anforderungen mit erfolgreichem Handeln zusammenhängen. Die Definition von Kompetenzen als kontextabhängig legt nahe, dass dabei sowohl Person- als auch Situationscharakteristika berücksichtigt werden müssen. Systematische Messungen setzen ein klares Modell der zu messenden Kompetenz voraus, das heißt, klare Vorstellungen darüber, in welchen Situationen, beziehungsweise bei welchen Anforderungen sich inter- und intraindividuelle Unterschiede zeigen sollten. Erst ausgehend davon lassen sich Testinhalte ableiten. Aus der Definition von kompetentem Verhalten wiederum lässt sich erst ableiten, wie das beobachtete Verhalten in Bezug auf die Kompetenz zu bewerten ist (Klieme & Hartig, 2007). Damit ist das Testinstrument an sich zuerst einmal zweitrangig (Wilson, 2005). Tests sollten in erster Linie darüber valide konstruiert werden, dass empirisch nachgezeichnet werden kann, welche kognitive Operation während der Aufgabenbearbeitung abläuft.

Neben dieser theoretischen Anforderung an die Kompetenzmodellierung sind auch in psychometrischer Hinsicht spezifische Ansprüche an die Kompetenzmessung zu stellen:

Das Problem der Definition von Kompetenzkriterien über eine kriteriumsorientierte Testung ist in erster Linie ein Problem der Integration von qualitativer und quantitativer Messung (Rost, 2004). Angestrebt wird ein qualitativer Messwert, das heißt, die Beschreibung der Kompetenz, die als Kriterium erreicht wurde. Zur Verfügung steht jedoch lediglich ein quantitativer Messwert in Form des Abschneidens in einem Test. Damit steht die Konzeption eines möglichst aussagekräftigen psychometrischen Modells im Zentrum einer präzisen Kompetenzmessung. Vor diesem Hintergrund ist die Intention von Kompetenzmessungen nicht nur daran orientiert, entsprechend den gängigen Gütekriterien ein valides, reliables und objektives Instrumentarium zu erstellen. Vielmehr sollte das psychometrische Messmodell mehr leisten können, als lediglich Auswertungsroutinen zur Verfügung zu stellen. Den Referenzrahmen für Kompetenzmessungen bietet deshalb die Absicht möglichst detaillierte Ergebnisse zu erhalten. Bei der Frage, welche Konstruktionsmodelle detaillierte Messinformationen über Kompetenzen erlauben, stehen zwei Möglichkeiten im Mittelpunkt, die sich idealerweise ergänzen (Klieme, Maag-Merki & Hartig, 2007; zur kritischen Diskussion auch Rost, 2004): Kompetenzniveaumodelle sowie Kompetenzstrukturmodelle. Koeppen, Hartig, Klieme & Leutner (2008) ergänzen Modelle die die Kompetenzentwicklung mitberücksichtigen, verweisen allerdings darauf, dass bis heute nur wenige Modelle diesen Aspekt mit ein beziehen, diese empirisch wenig abgesichert sind und zudem in ihren Entwicklungslinien unterschiedlich sind (bspw. Schnotz, Vosniadou & Carretero, 1999).

Kompetenzniveaumodelle erlauben die konkrete inhaltliche Beschreibung der erfassten Kompetenzen. Auf Basis der exakten Beschreibung der Anforderungen, der Aufgabenstellung und der Rahmenbedingungen können Kompetenzniveaus modelliert werden, auf Basis der Bewältigung der Anforderungen wiederum wird ersichtlich, welche Anforderungen auf welchem Kompetenzniveau bewältigt werden können. Neben der Generierung rein quantitativer Testwerte erlauben entsprechende

Kompetenzniveaumodelle eine kriteriumsorientierte, inhaltliche Auswertung der quantitativen Werte (Hartig & Klieme, 2006; auch Klieme, Maag-Merki & Hartig, 2007). Das einfachste Kompetenzmodell besteht darin, in einer Domäne verschiedene Stufen oder Level der Leistung zu definieren, die sich in der Elaboriertheit der kognitiven Prozesse unterscheiden (Rost, 2004). Voraussetzung für die Kriteriensetzung ist der Nachweis von Schwierigkeitsunterschieden zwischen den Aufgaben. An dieser Stelle wird nicht detaillierter auf die Konstruktion von Niveaumodellen eingegangen (für einen Überblick Hartig & Jude, 2007; Klieme & Hartig, 2007; Klieme, Maag-Merki & Hartig, 2007). Allerdings bleibt festzuhalten, dass die Qualität der Niveaufinition letztlich von der Qualität des theoretischen Modells abhängt, aus dem die Merkmale abgeleitet werden. Um der Beliebigkeit der Festlegung von Kompetenzstufen zu entgehen, können Methoden der Item-Response-Theorie zu Hilfe genommen werden, um sowohl Kompetenzen als auch Aufgabenschwierigkeiten abbilden zu können. Letztlich sind kriteriumsorientierte Kompetenzniveaumessungen dann von besonderem Interesse, wenn es darum geht, den Erfolg bestimmter Maßnahmen zu messen und Vergleiche zwischen dem beobachteten und dem idealiter zu erreichenden Niveau zu ziehen.

Ein weiterer Ansatz der Spezifizierung von Kompetenzen besteht darin, strukturbezogene Teilkompetenzen an bestimmte Merkmale von Testaufgaben zu knüpfen und den Zusammenhang spezifischer Aufgabenmerkmale mit entsprechenden Teilkompetenzen bei der Datenauswertung empirisch zu prüfen. Löst eine Person Aufgaben, die eine kognitive Teilkompetenz erfordern, andere Aufgaben basierend auf einer anderen kognitiven Kompetenz jedoch nicht, so resultiert daraus eine qualitative Aussage über die Leistungsstruktur (Rost, 2004). Das Modell kognitiver Teilkompetenzen ist den Kompetenzstrukturmodellen zuzuordnen: Hier steht vor allem die Frage nach der Dimensionalität von Kompetenzen im Vordergrund. Auch dieser Ansatz stellt einen Weg dar, statt eines rein quantitativen Urteils über die Erreichung eines Kompetenzstandards die Art der zu bewältigenden kognitiven Anforderungen und damit qualitative Unterschiede zu berücksichtigen. Ausgangsbasis dieser Modelle sind Annahmen über strukturelle Zusammenhänge zwischen unterschiedlichen Variablen. Dabei können nicht nur Zusammenhänge zwischen Kompetenzen verschiedener Bereiche (z.B. kognitive, soziale, motivationale, volitionale Dimensionen der Kompetenz und Handlungskompetenz) im Rahmen von Strukturmodellen interessant sein, sondern auch die Abbildung der Binnenstruktur bestimmter Kompetenzen. Wie in der Intelligenz- oder Persönlichkeitsdiagnostik allgegenwärtig, kann dementsprechend auch in der Kompetenzdiagnostik über faktorenanalytische Methoden die Dimensionalität von (Handlungs-)Kompetenz erfasst und interpretiert werden. Dabei basiert das Messmodell entweder bereits auf theoretischen Annahmen über Strukturen und diese werden anhand des psychometrischen Modells empirisch überprüft. Oder anhand eines mehrdimensionalen psychometrischen Modells wird empirisch von zugrundeliegenden Dimensionen auf latente Variablen geschlossen. Die Untersuchung von Zusammenhangsstrukturen ist letztlich nicht nur wertvoll, um Kompetenzen inhaltlich zu differenzieren und konkreter zu beschreiben, sondern auch, um sowohl das Zustandekommen von Kompetenzen als auch mögliche

Förderansätze zu untersuchen (für einen Überblick Hartig & Klieme, 2006; Klieme, Maag-Merki & Hartig, 2007).

Eine Charakterisierung möglicher Testaufgaben anhand ihrer Merkmale, sei es für die Bestimmung von Kompetenzniveaus oder Kompetenzstrukturen, bietet neben einer präzisen Auswertung demzufolge bereits für ihre Konstruktion notwendige Ankerpunkte. Die Aufgabenmerkmale bieten ein Raster, innerhalb dessen der Spielraum für die Aufgabenschwierigkeiten präziser definiert werden kann (Rost, 2004). Mit Blick auf eine Kompetenzmessung bleibt jedoch auch festzuhalten, dass es zu kurz greifen würde, den Fokus lediglich auf eine Feststellung der Ausprägung von Kompetenzen zu legen. Vielmehr gilt es, eine möglichst valide Kompetenzmessung darüber herzustellen, dass erst ein (1) theoretisch gründlich erarbeitetes und empirisch überprüftes Kompetenzmodell die Grundlage bietet für ein Testmodell, das sowohl (2) Binnenstrukturen, das heißt Teilstrukturen der Kompetenz, als auch (3) unterschiedliche Kompetenzniveaus sowie (4) Veränderungen im Lernen und im Prozess der Kompetenzentwicklung mit berücksichtigt (Koeppen, Hartig, Klieme & Leutner, 2008). Dieser Blick entspricht letztlich auch dem, was Vonken (2005) in Bezug auf das Kompetenzkonstrukt an sich festhält: Bei Kompetenz geht es längst nicht mehr nur um Kompetenz im Sinne einer allgemeinen situationsbezogenen Handlungsbefähigung, sondern um verschiedene Teilkompetenzen, die erst in ihrer Summe den Einzelnen zum Handeln und zur Bewältigung von Handlungszielen befähigen.

4. Zusammenfassung

Es wurde gezeigt, dass trotz fehlender Systematik im Umgang mit dem Kompetenzbegriff Weinert (2001) eine pragmatische Lösung anbietet. Über eine Differenzierung der Konstrukte Kompetenz und Handlungskompetenz schafft er einen Ansatzpunkt für eine funktionale Beschäftigung mit dem komplexen Konstrukt. Kompetenz als rein kognitives Konstrukt zu verstehen führt dabei nicht zu einer reduktionistischen Sichtweise, sondern erlaubt vielmehr eine funktionale Beschreibung auf theoretischer Ebene sowie eine pragmatische Analyse auf empirischer Ebene. Letztlich eröffnet gerade auch eine kognitionspsychologisch orientierte Kompetenzanalyse bedeutende Ansatzpunkte für eine Kompetenzförderung (auch Mayer, 2010): Gerade der Fokus auf eine Beschreibung individueller Unterschiede, auf Prozesse der kognitiven Kompetenzentwicklung sowie eine Analyse der Kompetenzstrukturen erlaubt eine detaillierte Sicht auf Lernen und Instruktion (auch Wittrock, 1989). Kompetenzniveau- bzw. Strukturmodelle bieten sich als messtheoretische Ansatzpunkte an. Voraussetzung ist eine hinreichend präzise theoretische Modellierung. Es muss dargelegt werden, in welcher Situation sich inter- und intraindividuelle Unterschiede zeigen. Dementsprechend wird im Folgenden das Konstrukt fächerübergreifender Problemlösekompetenz in seinen beiden Kernaspekten „Problemlösen“ und „fächerübergreifend“ theoretisch beschrieben.

Teil II: Theoretische Verortung fächerübergreifender Problemlösekompetenzen

5. Fächerübergreifendes Problemlösen: Theorien, Positionen, Befunde

Fächerübergreifende Problemlösekompetenzen werden definiert als situations- und inhaltsunabhängige kognitive Fähigkeiten, die bei der Bewältigung komplexer Problemstellungen abgerufen werden (Baumert, Klieme, Neubrand et. al., 2000; auch Fleischer, et. al., 2010; Mayer, 1994). Sie kommen zum Tragen, wenn neuartige, nicht explizit im Curriculum enthaltene und realistische Aufgabenstellungen unter Rückgriff auf vorhandene Problemlösekompetenzen sowie vorhandenes Wissen gelöst werden müssen. Fächerübergreifende Problemlösekompetenzen werden in unterschiedlichen Fächern und Disziplinen relevant und werden von disziplinspezifischem Wissen und disziplinspezifischen Kompetenzen unterschieden. Wie kann ausgehend von dieser Definition fächerübergreifendes Problemlösen theoretisch beschrieben und empirisch erfasst werden?

Die Definition zeigt bereits, dass die Problematik der Fassung des Konstrukts vor allem in seiner fächerübergreifenden Verortung liegt. Diese Verortung ist es, die das das Konstrukt für die Operationalisierung problematisch, für die Praxis jedoch so interessant macht, da sie die Überwindung der Trägheit von Wissen und die Frage nach disziplinübergreifender Anwendung fachspezifisch erlernter Problemlösekompetenzen fokussiert. Die Verortung fragt dabei auch nach der Situations- und Inhaltsunabhängigkeit des Konstrukts: Wie situations- und inhaltsunabhängig sind fächerübergreifend anzuwendende kognitive Problemlösekompetenzen, bzw. Teilkomponenten des Konstrukts?

Die folgende theoretische Modellierung beleuchtet deshalb neben der Operationalisierung der kognitiven Kompetenzen des Problemlösens insbesondere auch das Zusammenspiel zwischen disziplinbasiertem Erwerb kognitiver Problemlösekompetenzen und ihrer kontextübergreifenden Anwendung: Dargestellt wird zunächst, welche für die vorliegende Arbeit wesentlichen Befunde zum Prozess des Problemlösens bereits vorliegen und über welche Teilfertigkeiten der Prozess des Problemlösens beobachtet werden kann. In diesem Zusammenhang wird erläutert, dass kognitive Problemlösekompetenzen erst dann fassbar werden, wenn davon ausgegangen wird, dass das für den Prozess des Problemlösens konstitutive Verstehen der Problemsituation ebenso wie die schrittweise Veränderung auf planendem und schlussfolgerndem Denken basiert (Baumert, Klieme, Neubrand et. al., 2000; Mayer & Wittrock 1996; Sternberg & Davidson, 1992). Problemlösen wird als kognitive Verarbeitung der jeweiligen problemkonstituierenden Aussagen, bzw. Prämissen verstanden. Für die Modellierung kognitiver Kompetenzen beim fächerübergreifenden Problemlösen bedeutet dies, dass die jeweiligen Erscheinungsformen schlussfolgernden Denkens sowie mögliche Fehlleistungen, die zu fehlerhaften Schlussfolgerungen beim Problemlösen führen können, dargestellt werden müssen.

Dass fächerübergreifendes Problemlösen als Bewältigung neuartiger, nicht explizit im Curriculum enthaltener und realistischer Problemstellungen definiert wird, wirft weiterhin die Frage auf, wie fach(un)spezifisch Problemlösekompetenzen sind. Es wird deshalb die Relevanz des Kontextes

thematisiert sowie die Frage, wie der Transfer fachspezifisch erworbener Problemlösekompetenzen auf fächerübergreifende Problemstellungen gelingen kann.

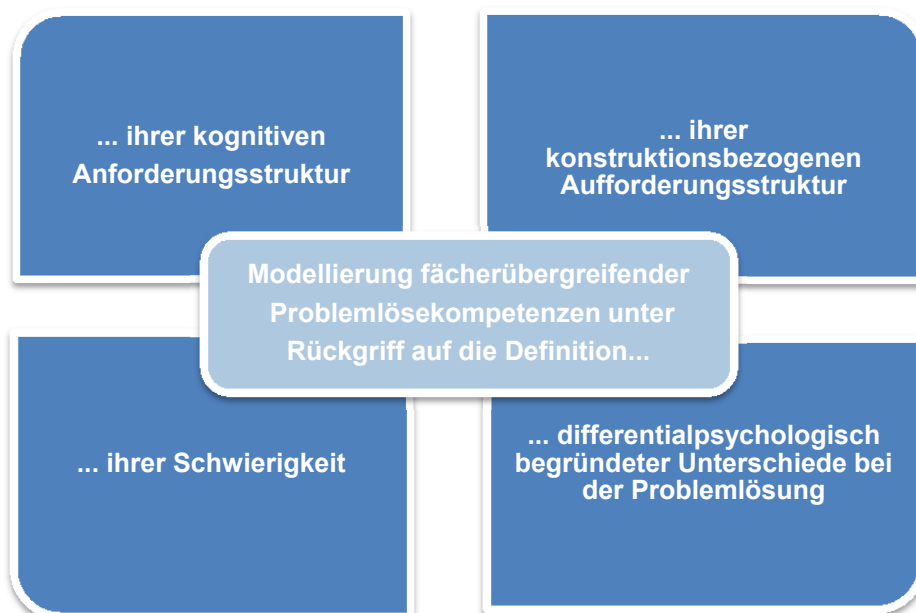
Eng damit verbunden sind auch Fragen der Charakterisierung von Problemstellungen, das heißt, problemimmanente Ansatzpunkte zur Manipulation der Schwierigkeit.

Daran anschließend wird auf personbezogene Kovariaten Bezug genommen, um Problemlösekompetenzen differentialpsychologisch zu betrachten. In diesem Zusammenhang wird die Rolle der Expertise dargestellt, bzw. die Frage danach, woran Unterschiede zwischen effektiven und weniger effektiven Problemlösern festzumachen sind.

Eine Darstellung bereits vorhandener Messinstrumente zur Analyse fächerübergreifender Problemlösekompetenzen schließt sich an.

Das dieser Arbeit zugrundeliegende Modell zur Konzeption fächerübergreifender Problemstellungen und Beschreibung fächerübergreifender Problemlösekompetenzen geht dementsprechend davon aus, dass fächerübergreifende Probleme sowohl über eine bestimmte konzeptionelle als auch kognitive Anforderungsstruktur modelliert werden können und die Schwierigkeit fächerübergreifender Probleme über problemimmanente sowie in der Person des Problemlösers liegende Ansatzpunkte manipuliert werden kann. Dieses Modell wird im Folgenden theoretisch hergeleitet und auf Basis empirischer Befunde begründet.

Abb.1: Modell zur Beschreibung fächerübergreifender Problemlösekompetenzen



5.1 Problemlösen – Theorien und Modelle

“What kind of computer is man?”

Earl Hunt, 1971

In den Alltag ist der Begriff des Problems und der Problemlösung fest integriert. Für eine präzise Fassung und Arbeitsdefinition des Begriffs stellt sich jedoch die Frage, wie Probleme überhaupt entstehen. Wann haben wir es mit einem Problem zu tun, oder anders formuliert, was macht ein Problem zu einem Problem? Im Gegensatz zur Erklärung des Prozesses, wie Probleme gelöst werden, ist diese Frage relativ einfach zu beantworten: Probleme ergeben sich unweigerlich aus der Zielorientiertheit menschlichen Handelns (Funke, 2003; auch Novick & Bassok, 2005). Dabei stellt jedoch nicht die Zielsituation an sich das Problem dar, vielmehr ergibt sich das Problem aus der Diskrepanz zwischen Ausgangs- und Zielpunkt sowie aus der Absicht diese Diskrepanz zu überwinden (Mayer, 2010, 1996; Funke, 2003; Jonassen, 2000; Davidson, 2003; Novick & Bassok, 2005; Mayer, 1995, 1992b; Ericsson & Hastie, 1994; Duncker, 1945). Hinzu kommt, dass ein Problem nicht einer (Routine-)Aufgabe entspricht, das heißt, die zur Lösung relevanten Handlungsprozeduren sind nicht vorhanden (ebd.). Insofern ergeben sich Problemstellungen erst dadurch, dass das anvisierte Handlungsziel nicht unweigerlich auch erreicht werden kann, sei es da die Mittel zur Zielerreichung unklar oder nicht vorhanden sind oder Ausgangs- und/oder Zielpunkt unklar sind. Eine erschöpfende Darstellung der Problemlöseforschung kann an dieser Stelle nicht geleistet werden (für einen Überblick u.a. Mayer, 2010, 1979; Adler & Rips, 2008; Anderson, 2007; Bassok & Novick, 2005; Holyoak & Morisson, 2005; Davidson & Sternberg, 2003; Kotovsky, 2003; Ericsson & Hastie, 1994; Sternberg, 1994; Gilhooly, 1982; Posner, 1976). Im Folgenden soll ein Überblick lediglich auf diejenigen Aspekte verweisen, die für die vorliegende Arbeit von Bedeutung sind.

Zentrale Fragestellungen im Zusammenhang mit Problemlösen

Problemlösen, Denken, Erkennen, Entscheiden – vielfältige Begriffe werden im Zusammenhang mit Problemlösen oftmals synonym verwendet. Trotz der umfassenden Forschungstradition in Bezug auf Denken und Problemlösen liegt bislang kein Konsens in Bezug auf eine definitorische Abgrenzung vor (Mayer, 2010; Funke, 2003; Kotovsky, 2003; Ericsson & Hastie, 1994; Anderson, 1990). Ebenso wenig gibt es ein endgültiges Modell über die zugrundeliegenden kognitiven Prozesse und Mechanismen. Bereits im Hinblick auf die Frage, wie Problemlösen definitorisch gefasst werden kann, existieren unterschiedliche Vorschläge (bspw. Funke, 2003; Lür & Spada, 1990; Dörner, 1976; Klix, 1971; Duncker, 1945; Polya, 1945). Ein Vergleich zeigt, dass fast alle Definitionen die Bestandteile eines Problems festhalten: Ein Problem besteht aus einem Ausgangspunkt der inkohärent mit einem gewünschten Ziel ist. Weiterhin wird formuliert, dass ein Set von Operatoren zur Überwindung der Diskrepanz zwischen Ausgangs- und Zielpunkt bedeutsam ist. In Bezug auf eine präzisere Charakterisierung des Problemlösens sowie des Problemlöseprozesses ist der kleinste gemeinsame

Nenner jedoch lediglich der, den Prozess des Problemlösens als Suche zu definieren (Anderson, 2007; Funke, 2003; Chi, Glaser, & Rees, 1982). Wird der Prozess der Suche näher betrachtet, so zeigt sich, dass sie zu ihrer näheren Bestimmung als *gerichtet*, das heißt die Ausrichtung der Suche hin auf die Erreichung eines Zielzustandes sowie als *prozesshaft*, das heißt, das Lösen eines Problems über Manipulationen bzw. Operationen innerhalb des kognitiven Systems der Problemrepräsentation, definiert wird. Als drittes Element wird die Suche als *kognitiv* und internal ablaufend aufgefasst (Mayer, 2010, 1992, 1979; Anderson, 2007, 1990; Funke, 2003; Sternberg & Davidson, 2003; Jonassen, 2000).

In Bezug auf eine definitorische Fassung von Problemlösen kann dementsprechend festgehalten werden, dass problemlösenden Denkens ein beobachtbares Verhalten erzeugt, das die denkende Person über die Aktivierung des Wissens- und Operatorrepertoires von einem Ausgangs- in einen Zielzustand bringt (Mayer, 2010). Dies gestaltet sich als kognitiv ablaufende Suche, um eine Lücke in einem Handlungsplan zu füllen, der nicht routinemäßig zu bewerkstelligen ist (Funke, 2003). Die leitende Fragestellung im Zusammenhang mit Problemlösen ist dementsprechend nicht nur, wie Problemstellungen über die zentralen Charakteristika des Anfangs- und Zielzustandes hinaus konkreter beschrieben werden können, sondern vielmehr auch, wie der internale Prozess des Problemlösens, das heißt die zur Problemlösung relevanten kognitiven Fertigkeiten und Prozesse näher definiert werden können. Das heißt, auf welche Befunde kann zur Beschreibung der kognitiven Anforderungsstruktur und konstruktionsbezogenen Aufforderungsstruktur von Problemen zurückgegriffen werden?

Antwortmodelle: Theorien des Problemlösens

Was machen Personen, wenn sie über Probleme nachdenken und wie versuchen sie, Probleme zu lösen? Diese Fragen wurden bereits um die Jahrhundertwende des 19. Jahrhunderts gestellt (Holyoak & Morriison, 2005; auch Funke, 2003; Hunt, 1994; Greeno & Simon, 1988) und im Gegensatz zu den Untersuchungsmethoden blieben die leitenden Fragen bis heute relativ unverändert: Was kann über die Prozesse des problemlösenden Denkens ausgesagt und wie können internal ablaufende Prozesse zuverlässig untersucht werden? Eine Vielzahl an Antworten wurde darauf gegeben: Von Ansätzen wie der Problemlösung als erlerntes Reaktionsmuster (Behavioristen), über Problemlösen als Veränderung der „defekten Gestalt“ über Prozesse der Einsicht, Intuition und Aha-Effekte (Gestalttheoretiker) bis hin zu mathematischen Modellierungen kognitiver Problemlösesysteme (Informationsverarbeitungstheoretiker). Daneben stehen psychoanalytisch und evolutionsbiologisch ausgerichtete Annahmen zum problemlösenden Denken (für einen detaillierten Einblick bspw. Holyoak & Morriison, 2005; Novick & Bassok, 2005; Funke, 2003). Den bis heute einflussreichsten Ansatz bietet die Informationsverarbeitungstheorie, wobei es ihr Verdienst ist, den Schwerpunkt der Beschäftigung mit Problemlösen auf die kognitiven Prozesse zu konzentrieren und präzise Prozessmodelle

problemlösenden Denkens einzufordern. Dadurch, dass sie Problemlösen ebenfalls als Strukturierung und Umstrukturierung in Richtung der Zielerreichung auffassen, greifen ihre Annahmen auch wesentlich auf die Gestalttheorie zurück.

Im Folgenden werden die für das vorliegende Modell fächerübergreifenden Problemlösens relevanten Befunde herausgegriffen und in der Forschungshistorie verortet:

Als Vorläufer des Behaviorismus beschäftigte sich der Assoziationismus mit der Frage, wie geistige Tätigkeit und Problemlösen ablaufen könnte. Assoziationstheoretiker gehen davon aus, dass problemlösendes Denken als Bewegung von einer „Vorstellung“ zur nächsten über Assoziationsketten funktioniert (Gilhooly, 1982). Im Rahmen der behavioristischen Beschäftigung mit Problemlösen wird diese Annahme von Assoziationsketten weitergeführt, Denken wird als Absuchen von Reaktionshierarchien definiert (Thorndike, 1898; auch Mayer, 1979; Krech, Crutchfield & Livson, 1974). Problemlösen wird als Umstrukturierung von Reaktionshierarchien, das heißt als instrumentelles Lernen verstanden. Die Annahmen der Assoziationstheoretiker werden beim Konzept- und Regellernen (bspw. Hull, 1920 nach Gilhooly, 1982) oder auch bei der Frage nach Strategieranwendung und Hypothesentesten beim Problemlösen (bspw. Bruner, Goodnow & Austin, 1956) aufgenommen. Ihre Überlegungen spielen eine Rolle bei Lernprozessen (bspw. Anderson, 1993) oder bei Annahmen über Assoziationsstärken bei der kognitiven Modellierung (bspw. Strube, 1998).

Der Frage, wie Problemlösen als kognitiver Prozess weiter differenziert werden kann, widmeten sich die Gestalttheoretiker. Sie fragten nach internalen Prozessen während der Problemrepräsentation und Problemlösung (Funke, 2003) und untersuchten, wie Informationen im Zusammenhang mit Problemen gesehen, interpretiert und organisiert werden. Bei einem Problem als „defekter Gestalt“ steht für sie die Suche nach Schwierigkeiten bei der Problemlösung im Vordergrund, wobei Prozesse der Wahrnehmung, die Relevanz und Interferenz von Vorwissen eine Rolle spielen (Novick & Bassok, 2005). Wahrnehmungsnähe (bspw. Dunckers 9-Punkte-Problem oder Kerzenproblem) und Einstellung (bspw. Luchins Wasserumfüllaufgabe) beeinträchtigen die Problemlösung und können zu Fixierung und funktioneller Gebundenheit führen. Ein wichtiger Befund ist, dass die Vorverwendung eines Gegenstandes sowie die Abstraktheit des Problems seiner Lösung entgegenstehen können (bspw. Dunckers Kerzenaufgabe). Neben dem negativen Einfluss von Vorwissen fragen sie auch nach positiven Einflüssen und stellen fest, dass Hilfestellungen und Vorerfahrungen dann hilfreich sind, wenn sie zu einem verbesserten Problemverständnis und einer Neubewertung der Situation führen (bspw. Saugstadt & Raaheim, 1968; Maier, 1945). Die Gestalttheoretiker zeigen, dass Problemlösen einen Abgleich des Gegebenen mit dem Geforderten verlangt, wobei produktives und auf Strukturverständnis ausgerichtetes Denken statt reproduktives, mechanisches hilfreich ist (Mayer, 1979). Zudem fragen sie erstmals nach Phasen des Problemlöseprozesses (Posner, 1976). Sie unterscheiden zwischen *Preparation*, *Incubation*, *Illumination* und *Verification* (Wallas, 1926) und versuchen, diese Phasen empirisch zu fundieren (Duncker, 1945; Polya, 1945 nach 1985). Auch der

Lösungsprozess wird versucht zu strukturieren (bspw. Lösungsbäume bei Seilaufgabe oder Hobbits-und-Orks Aufgabe, Restle & Davis, 1962).

Relevante Befunde zu kognitiven Prozessen des Problemlösens können auch die Bedeutungstheoretiker berichten, ihre Befunde sind auch mit Blick auf die Frage des Transfers interessant. So fokussieren sie den Einfluss externer Aspekte, bspw. die Struktur der Darbietung des Problems oder die assoziierte Beziehung der Problemelemente. Kernannahme ist, dass die Umstrukturierung des Problems gesteuert wird über den Abgleich der Problemelemente mit im Gedächtnis abgespeicherten Begriffen (Solso, 2005). Dabei nimmt die Reihenfolge der Begriffsdarbietung Einfluss auf die Problemlösung, Schemata als organisierte Repräsentationen von Erfahrungen spielen eine zentrale Rolle, zwischen bedeutungsvollem und mechanischem Wissen muss unterschieden werden (bspw. Bransford & Johnson, 1972; Cofer 1967). Weiter zeigen Befunde, dass verständnisorientiertes, eigenaktives Lernen und die Möglichkeit Assimilationen zu erleichtern (bspw. über Vorinformationen) positiv auf Problemlöseerfolg und Transfer wirken im Gegensatz zu eher mechanischem Einüben (bspw. Bransford & Johnson, 1972; Gagné & Smith, 1962). In diesem Zusammenhang ist auch die Theorie der Semantischen Netze zu nennen, die Fragen der kognitiven Modellierung, der kognitiven Repräsentation von Problemen sowie der Rolle von Wissen und Gedächtnisstrukturen in den Blick nimmt (bspw. Neokonnektionismus, Strube, 1990). Die Kernannahme ist, dass die Struktur des Wissens und der Abruf von Gedächtnisinhalten Einfluss auf die Problemlösung nimmt. Zentraler Befund ist, dass das aktive Bemühen um Bedeutung die Problemlösung beeinflusst (Sweller 1994; Bartlett, 1932). Die Assimilation neuer Informationen hängt davon ab, welche Informationen bereits vorhanden sind. Treten Veränderungen auf, fallen bspw. Details weg oder werden in Übereinstimmung gebracht, um neue Informationen an Vorhandene anzupassen (Bartlett, 1932). Die Betonung des Zusammenhangs von Problemlösen und Gedächtnis ist für die Beschäftigung mit Problemlösen zentral, da der Blick auf Schemata, auf Fragen des Wiedererkennens, Erinnerns, Kategorisierens sowie den Prozess der Bedeutungs(re)konstruktion bei der Bearbeitung neuer Aufgaben gelenkt wird (Chi, Glaser & Rees, 1982).

Im Rahmen der Informationsverarbeitungstheorie schließlich wird Problemlösen mit Fokus auf eine funktionalistische Beschreibung des kognitiven Systems aufgegriffen (Ohlsson, 2012). Ihre Annahmen sind zwar modellhaft, idealtypisch und damit begrenzt anwendbar; ihre Intention, den Problemlöseprozess in Modelle zu übersetzen, führt jedoch zur Präzisierung zentraler Annahmen. Zudem werden Fragen nach der Problemrepräsentation und Fragen nach Problemlöseprozessen gemeinsam betrachtet (Novick & Bassok, 2005; Solso, 2005). Modelle hierarchischer Operationen werden definiert, um Prozessablauf und Strategien näher beschreiben zu können und Regularien für die Suche nach der Problemlösung zu finden (bspw. General Problem Solver, Newell & Simon, 1961; Logic Theorist, Newell & Simon, 1957, 1970). Erstmals werden empirisch überprüfbare methodische Ansätze für eine Theorie menschlicher Denkprozesse aufgestellt. Zur Beschreibung des Problemlösens schlagen Newell und Ernst (1969; auch Newell & Simon, 1970) die auch heute noch geltenden

Komponenten *Anfangszustand*, *Zielzustand*, *Operatoren* und *Problemraum* vor. Der Prozess des Problemlösens besteht darin, den richtigen Pfad durch den Problemraum – die Gesamtheit aller möglichen Problemzustände – zu entdecken. In Bezug auf die Lösungsgüte und -schnelligkeit spielen Vorerfahrungen eine Rolle (Bekanntheit der Algorithmen, Wissen, heuristische Vorerfahrungen, Faustregeln; bspw. Rumelhardt, Lindsay & Norman, 1972). Die Rolle von Kurzzeit-, Arbeits- und Langzeitgedächtnis wird integriert und die Frage nach Expertise gestellt, weitere personbezogene Faktoren jedoch nicht berücksichtigt (Greeno, 1973). Das Ziel ist, Invarianten im kognitiven Prozess des Problemlösens zu finden und damit den Prozess des Problemlösens vollständig zu erklären (Simon, 1990). Ihr bedeutendster Beitrag liegt allerdings darin, Heuristiken der Problemlösung aufzuzeigen (Ohlsson, 2012; Anderson, 1987).

Die Weiterentwicklung der Problemlöseforschung konzentriert sich in der Folge zunehmend auf die Eigenschaften von Problemen (Variablenanzahl, Vernetztheit, Transparenz, Eigendynamik, Art der Barriere, Polytelie; bspw. Dörner, 1976): Eine Problemstellung zeichnet sich über bestimmte Eigenschaften aus und ist definiert über eine seine inhärente Struktur und die Beziehung der Problemelemente untereinander. Auch die Merkmale von Operatoren werden nun mitberücksichtigt (Wirkungsbreite, Reversibilität, Anwendungsvoraussetzungen, Wirkungssicherheit). Im Hinblick auf den Einsatz der Operatoren wird bspw. das Wissen der Problemlöser um die Eigenschaften eines Operatorsystems berücksichtigt. Daneben dominieren Fragen der Relevanz höherer kognitiver Prozesse im Zusammenhang mit Problemlösen (Ohlsson, 2012). Auf diese wird nun im Folgenden eingegangen, um konkrete kognitive Anforderungen im Zusammenhang mit Problemlösen zu beschreiben.

Zusammenfassung

Problemlösen ist zielorientiertes Denken und Handeln in Situationen, für deren Bewältigung keine Routinen verfügbar sind. Es liegt ein mehr oder weniger gut definiertes Ziel vor, der Problemlöser weiß aber nicht unmittelbar, wie es zu erreichen ist. Diese Inkongruenz von Zielen und verfügbaren Operatoren ist konstitutiv für ein Problem (Mayer & Wittrock 1996; Sternberg & Davidson, 1992). Voraussetzung für die Problemlösung ist das Verstehen der Problemsituation sowie die schrittweise Veränderung der Problemelemente, um zu einer Lösung zu gelangen. Für die Konzeptionalisierung der kognitiven Anforderungsstruktur stellt sich im Weiteren allerdings die Frage, vor welche kognitiven Anforderungen Probleme stellen.

5.2 Kognitive Anforderungen beim Problemlösen

Problemlösen wird als kognitiver Prozess innerhalb sequenzieller Problemlösephasen definiert (Brophy, 1998; Mumford, Reiter-Palmon & Redmond 1994; Bransford & Stein, 1993; für einen

Überblick Pretz, Naples & Sternberg, 2003). Die Phase des Problemverständnisses, die Problemanalyse, stellt die bedeutendste Phase dar. Zur Beschreibung aller Problemlösephasen kann auf unterschiedliche Modelle zurückgegriffen werden (bspw. Polya, 1945), wobei die Problemlösephasen wie sie im Rahmen des IDEAL-Modell (Bransford, Sherwood & Strodbeck, 1951) beschrieben sind, trotz der Diversität der theoretischen Ansätze zur Beschreibung des Problemlöseprozesses als Konsens angenommen werden können:

1. Bestimmung des Zieles (**I**dentify)
2. Analyse der Ausgangssituation und Aufbau einer mentalen Repräsentation, eines Situationsmodells (**D**efine)
3. Bestimmung der Lösungsstrategie und Planung von Lösungsschritten (**E**valuate)
4. Ausführen des Lösungsplans, begleitende Kontrolle und ggfs. Modifizierung der Lösung (**A**pply)
5. Evaluation der Lösung (**L**ook back)

Lipshitz, Levy und Orchen (2006) konnten zeigen, dass diese Phasen bei Problemlösern als implizites kognitives Schema vorliegen. Weiter konnten sie zeigen, dass Problemlöser umso effektiver sind, je mehr Wert auf den Prozess der Problemanalyse gelegt wird.

Problematisch ist jedoch, dass von den beschriebenen Problemlösephasen nicht direkt auf das beobachtbare Verhalten geschlossen werden kann, die Problemlösephasen besitzen lediglich heuristischen Wert und dienen als Kategoriensystem. Fassbar werden sie erst, wenn davon ausgegangen wird, dass das für den Prozess des Problemlösens konstitutive Verstehen ebenso wie die schrittweise Veränderung auf planendem und schlussfolgerndem Denken basiert (Baumert, Klieme, Neubrand et. al., 2000; Mayer & Wittrock 1996; Sternberg & Davidson, 1992). Die Phasen bieten demzufolge Anhaltspunkte zur Ableitung derjenigen kognitiven Kompetenzen die bei der Bewältigung dieser Phasen relevant werden.

Problemlösen als kognitive Verarbeitung von Aussagen

„[W]henever one cannot go from the given situation to the desired situation simply by action, then there has to be recourse to thinking.“ (Duncker, 1945, S. 1)

Denken schafft die Voraussetzung zum planenden Handeln. Inwieweit Denken und Problemlösen zu differenzieren sind, ist bis heute unklar und Gegenstand unterschiedlicher Definitionen (bspw. Anderson, 1985, 2007; Funke, 2003, 2006). Festzuhalten ist, dass beim Problemlösen über Denken Lücken geschlossen werden. Problemlösungen verlangen dabei einen Vergleich von Aussagen, um zu einem gültigen Schluss zu gelangen (Anderson, 2007). Denken im Rahmen der Problemlösung basiert demzufolge auf der widerspruchsfreien und folgerichtigen schlussfolgernden Verknüpfung der

jeweiligen problemkonstituierenden Elemente. Schlussfolgerndes Denken vollzieht sich in unterschiedlichen Formen (Funke, 2006):

- als logisches Schließen, bei dem deduktive Urteile getroffen werden,
- als Wahrscheinlichkeitsurteil, bei dem induktive Schlüsse auf Basis unvollständiger Informationen gefällt werden oder
- als kausales Schließen, einer Sonderform des induktiven Schließens.

Wie kann schlussfolgerndes Denken beschrieben werden? Schlussfolgerndes Denken bezieht sich auf den Übergang von einer oder mehreren Aussagen, bzw. Prämissen, zu einer neuen Aussage, bzw. Konklusion (Kutschera & Breitkopf, 1985). Schlussfolgerungen beruhen darauf, eine Aufgabe in ihre konstituierenden Elemente zu zerlegen, um im Abgleich aller gegebenen Elemente über logische Verknüpfungen zu einer Lösung zu gelangen. Relevant werden somit Prozesse des Enkodierens und des logischen Verknüpfens gegebener Informationen. Aus den Problemelementen und ihrer Inbezugsetzung werden Regeln abgeleitet, um zu Schlussfolgerungen zu gelangen, die auf den vorliegenden Sachverhalt angewendet werden können.

Dabei wird die Überprüfung der Informationen oder auch Prämissen auf ihre Gültigkeit und logische Konsequenz hin als Syllogismus bezeichnet (Anderson, 2007; Solso, 2005; Zimbardo & Gerrig, 2004; Mayer, 1992):

Alle Menschen sind sterblich. (Hauptprämisse)

Sokrates ist ein Mensch. (Nebenprämisse)

Auch Sokrates ist sterblich. (Schlussfolgerung)

Syllogismen bestehen dementsprechend aus Prämissen und einer Schlussfolgerung, wobei Prämissen verschiedene Formen annehmen können. Es werden vier Haupttypen unterschieden (Anderson, 2007; Evans, 2005): Alle A sind B (universell bejahender Satztypus), keine A sind B (universell verneinender Satztypus), einige A sind B (speziell bejahend), einige A sind nicht B (speziell verneinend). Prämissen bestehen aus einem Antecedens (Wenn-Teil) und einem Konsequens (Dann-Teil). Die Schlussregeln entsprechen entweder einem Modus ponens, der die Ableitung der Konsequenz einer bedingten Aussage erlaubt, wenn das Antecedens gegeben ist:

Wenn A, dann B

A ist gegeben

Es folgt also B (Modus ponens)

Oder einem Modus tollens, der die Negation des Antecedens aus der Negation des Konsequens schließt:

Wenn A, dann B

B ist falsch oder nicht gegeben

A ist falsch oder nicht gegeben (Modus tollens)

Jonsohn-Laird (1995) unterscheidet weiter syllogistische Schlussfolgerungen (Quantoren wie bspw. alle, keine, wenige), konditionale Schlussfolgerungen (Prämissen durch Junktoren wie und, oder, wenn, dann verbunden) und relationale Schlussfolgerungen (Beurteilung von Verhältnissen wie bspw. größer als, teurer als muss mitberücksichtigt werden). Logische Schlüsse können in zwei verschiedenen Erscheinungsformen vorliegen: Dem induktiven Schluss, als dem Prozess des Schließens vom Besonderen auf das Allgemeine, der einer Wahrscheinlichkeitsaussage entspricht sowie dem deduktiven Schluss, das heißt, dem Ziehen von Schlüssen aus allgemeinen Prinzipien von denen angenommen wird, dass sie wahr sind. Modelle des schlussfolgernden Denkens gehen davon aus, dass Schlussfolgern entweder auf Basis der Merkmale der gegebenen Prämissen, die miteinander verglichen werden, abläuft oder über eine lineare Ordnungsbildung (Evans, 2005; Rips, 1994). Syllogismen können gelöst werden über kognitive Strategien wie Vorwärtsverkettungen (M-P, S-M \rightarrow S-P), Reizketten (P-M, S-M \rightarrow S-P), Reaktionsäquivalenzen (M-P, M-S \rightarrow S-P) oder auch Rückwärtsverkettungen (P-M, M-S \rightarrow S-P) (Anderson, 2007; Evans, 2005; Solso, 2005; Zimbardo & Gerrig, 2004). Übersetzt auf die Problemstruktur bedeutet dies, dass über ihre Zerlegung in Teilkomponenten Probleme in die Struktur von Syllogismen übersetzt werden, im Rahmen derer die logischen Beziehungen zwischen den Prämissen, das heißt den Einzelinformationen des Problems erschlossen werden (Rips, 1994; Greeno & Simon, 1988). Problemlösungen verlangen dementsprechend einen Vergleich ihrer jeweiligen Aussagen, um zu einem gültigen, logisch richtigen Schluss zu gelangen. Welche Fehler können nun beim logischen Schließen auftreten?

Fehlerquellen im logischen Schließen

Biases werden auf verschiedene Ursachen zurückgeführt: Sie kommen bspw. zustande, da Problemlöser problemrelevante Information bereits vor der analytischen Schlussfolgerungsphase vernachlässigen, oder zwar analytisch vorgehen, aber die adäquaten mentalen Modelle für die Prämissen nicht bilden (Beller, KuhnMünch & Spada, o. J.). Oftmals werden Regeln fehlerhaft angewendet. Braine (nach Knauff, 2006) unterscheidet diesbezüglich Verständnisfehler (Prämissen werden falsch verstanden), Strategiefehler (eine direkte Lösung ist mit den vorhandenen Schlussschemata nicht möglich) sowie Prozessfehler (die Kapazität des Arbeitsgedächtnisses ist

überlastet). Fehlerquellen können auch darin liegen, dass die formale Logik von der mentalen Logik des Problemlösers auf Basis des Alltagswissens abweicht. Im Folgenden wird auf diejenigen Fehlerquellen näher eingegangen, die in der Konstruktion des logischen Schlusses begründet sind (für eine Zusammenfassung weiterer Fehlerquellen und -theorien bspw. Knauff, 2006; Rips, 1994): Die Form von Prämissen, deren Inhalt und die Interferenz des Inhaltes mit Vorwissen und Erfahrungen der Problemlöser sowie die Komplexität der Prämissen.

Die Annahmen von Fehlern aufgrund der Form der Prämissen gehen von zwei verursachenden Aspekten aus: Einerseits verleitet die Atmosphäre oder Stimmung, die durch die Prämissen entwickelt wird zu einem logisch falschen Schluss, andererseits wird angenommen, dass logisch falsche Schlüsse durch Umkehrungseffekte begünstigt werden.

Der Fehlertypus der Atmosphäre wird erklärt über Umgebungseffekte, das heißt über die Rahmung der Prämissen (Hawkins, Glick, Pea & Scribner, 1984; Woodward & Sells, 1935). Die Form der Prämisse begünstigt eine Schlussfolgerung, die nach der gleichen formalen Gegebenheit konstruiert ist. Dieser Atmosphären-effekt bewirkt, dass eine Schlussfolgerung aufgrund seiner Form akzeptiert/nicht akzeptiert wird. So schaffen bejahende Prämissen günstige Voraussetzungen für bejahende Schlussfolgerungen, verneinende Prämissen für verneinende Schlussfolgerungen, auf Quantoren beruhende Prämissen schaffen Voraussetzungen für ebenfalls auf einer Quantifizierung beruhende Schlussfolgerungen. Chapman und Chapman (1959) konnten dies ebenso belegen wie Johnson-Laird und Byrne (1989) oder auch Wilkes (1928), der zeigte, dass aus einer Auswahl unterschiedlicher Schlusssätze derjenige Schlusssatz überzufällig häufig gewählt wurde, der der Atmosphäre der Prämissen entspricht:

Einige Republikaner haben Ölquellen erworben.

Einige Wallonen haben Ölquellen erworben

Einige Wallonen sind Republikaner

(Wilkes, 1928)

Es wird davon ausgegangen, dass Umgebungseffekte sowohl den kognitiven Prozess der linearen Ordnungsbildung beeinflussen, als auch die Möglichkeit, die gegebenen Aussagen in Form einer Art räumlicher Logik zu repräsentieren (DeSoto nach Mayer, 1979; für einen Überblick Knauff, 2006). Ausgehend davon zeigte sich im Umkehrschluss auch, dass Leistungen im logischen Schließen dann besser waren, wenn eine bestimmte Aussagenfolge eingehalten wurde, bzw. wenn sich die zweite Prämisse gut in die von der ersten Prämisse definierte Beziehung einfügt (Huttenlocher, 1986). Im Zusammenhang damit ist anzumerken, dass auch die Bildhaftigkeit eines Syllogismus Einfluss nimmt (Falmange & Gonsalves, 1995; Johnson-Laird, 1995; Clement & Falmagne, 1986). Befunde zeigen,

dass sehr bildhafte Aussagen signifikant besser gelöst werden als abstrakte Aussagen – hier wurde das Erstellen einer Repräsentation erschwert (auch Duncker, 1945). Im Hinblick auf die Art der Repräsentation halten Knauff (2006) und Johnson-Laird (1995) jedoch fest, dass für die Leistung beim Schlussfolgern letztlich die Einfachheit, mit der eine räumliche Repräsentation (bspw. „über“, „unter“) konstruiert werden kann, entscheidend ist, wobei der Konstruktionsprozess durch die Art der Vorstellung beeinträchtigt werden kann. So können rein visuelle Relationen (bspw. „schmutziger“, „sauberer“) den Aufbau einer Repräsentation sogar behindern (*visual-impedance effect*, Knauff, 2006), wohingegen rein räumliche oder zusätzlich räumliche Relationen den Aufbau mentaler Repräsentationen begünstigen können. Kontrollrelationen (bspw. „schlechter“, „besser“) sind sowohl räumlich als auch visuell schwierig vorstellbar.

Im Hinblick auf Fehler aufgrund von Umkehrungseffekten wird schließlich angenommen, dass die Form des Syllogismus nach dem Muster AB – BC oftmals den falschen Schluss AC nach sich zieht (Lefford, 1946):

Einige Philosophen sind Menschen (AB)

Einige Menschen sind fehlbar (BC)

Philosophen sind fehlbar (AC)

(Lefford, 1946)

Fehlerquellen sind daneben auch Erfahrungen und Vorwissen:

In zahlreichen Studien wurde belegt, dass logisch falsche Schlüsse dadurch entstehen können, dass Schlussfolgerungen, die eher mit unserer Erfahrung übereinstimmen, favorisiert werden, wie das nachfolgende Beispiel verdeutlicht (für einen Überblick Stanovich, 2004): Sind Katzen Hunde?

Alle Katzen benötigen Wasser

Hunde benötigen Wasser

Alle Katzen sind Hunde

(Stanovich, 2004)

Dieser Schluss wird unmissverständlich als logisch falscher Schluss empfunden. Warum ist dies so? Mit (in)konsistenten oder auch (un)erwünschten Prämissen wird oftmals so verfahren, dass kognitive Konsistenz erreicht oder erhalten bleibt, auch auf Kosten der Einhaltung logischer Regeln. Gefolgt wird der Tendenz nach Widerspruchslosigkeit der Information (Nickerson, 1998). Eine große Rolle spielt die Nähe und Bekanntheit eines Sachverhaltes:

Alles was einen Motor hat, braucht Öl
Autos brauchen Öl

Alle Autos haben Motoren
(Solso, 2005)

Dieser Schluss wird als korrekt erachtet, da bekannt ist, dass alle Autos Motoren haben. Aus logischer Sicht ist er jedoch falsch. Problematisch werden auf diese Art gezogene Schlussfolgerungen vor allem dann, wenn aufgrund von Halbwissen Schlüsse gezogen werden und Übergeneralisierungen die Folge sind, um Aussagen in bestehende kognitive Strukturen einzupassen. Die Fehler treten nicht bei unbekanntem Kontexten auf (Markowitz & Nantel, 1989), inhaltlich nicht gebundene Syllogismen werden zudem schneller gelöst, da die kontextuelle Bindung nicht interferiert (Parrott, 1949, nach Mayer, 1979; McGuire, 1960).

Daneben nimmt Erfahrung über einen weiteren Aspekt Einfluss, wie das nachfolgende Beispiel zeigt:

Alle Menschen sind moralisch
Putin ist ein Mensch

Putin ist moralisch
(Solso, 2005)

Die gegebenen Prämissen müssten den logisch korrekten Schlusssatz nach sich ziehen, dass Putin moralisch ist. Dieser logisch korrekte Schluss fällt jedoch aufgrund unserer Kenntnis von Putin als einem unmoralischen Menschen schwer. Material mit stark emotional gefärbtem Inhalt zieht mehr Fehler nach sich, als Material mit neutralen Informationen (Nickerson, 1998).

Wie kann dies erklärt werden? Die Begründung liegt darin, dass Erfahrungen ausschlaggebend dafür sind, wie mit vorliegenden Aussagen umgegangen wird (Goldstone & Son, 2005, Tversky, 1977, Tversky & Kahneman, 1974):

„From causes which appear similar we expect similar effects“ (Goldstone & Son, nach Sloman & Lagnado, 2005, S. 95).

Das heißt, vorliegende Sachlagen werden mit bereits vorhandenen Erfahrungen abgeglichen, es werden Voraussagen getroffen und logische Schlüsse abgeleitet. Dieser Abgleich wird selbst dann vorgenommen, wenn zwar eine Vorerfahrung über eine Verknüpfung bestimmter Variablen besteht, diese beiden Variablen im vorliegenden Sachverhalt jedoch völlig unverknüpft sind (Stanovich, 2004).

Vor allem induktive Schlüsse basieren auf einem Abgleich von Ähnlichkeitsbeziehungen zwischen Merkmalen vorhandener Erfahrungen und Schemata sowie vorliegender Sachlagen. Die Lücke in unklaren Situationen wird gefüllt und es wird auf dieser Basis eine Wahrscheinlichkeitsaussage getroffen. Ähnliche und unterschiedliche Merkmale in Prämissen sind dabei ausschlaggebend für die Zuschreibung von Ähnlichkeitsbeziehungen (Buehner & Cheng, 2005; Goldstone & Son, 2005). Ob Merkmale als ähnlich eingeschätzt werden, orientiert sich an einem Ähnlichkeitsabgleich über typische Anhaltspunkte und Ankerheuristiken (Tversky & Kahneman, 1974). Je mehr Ähnlichkeiten zwischen Erfahrung und vorliegendem Sachverhalt gefunden werden, umso eher werden mögliche weitere Erklärungen zugunsten einer Schlussfolgerung im Einklang mit bisherigen Erfahrungen übersehen (Evans, 2001; Goldstone & Son, 2005). Je weniger sich Merkmale ähneln, umso stärker leitet die gegebene Prämisse. Allerdings führen Vorerfahrungen nicht ausschließlich zu ineffektiven logischen Schlüssen, sondern können auch den Einsatz erfolgreicher und effektiver Strategien nach sich ziehen (bspw. Wason-Aufgabe, Sloman & Lagnado, 2005; Cheng & Holyoak, 1985; Johnson-Laird & Wason, 1970). Zudem werden mit zunehmender Expertise weniger Oberflächenbeziehungen und mehr Tiefenbeziehungen zum Vergleich herangezogen (Suzuki, Ohnishi & Shigematsu, 1992).

Schließlich ist im Hinblick auf Einflussgrößen anzumerken, dass auch mit zunehmender Anzahl der Items Schwierigkeiten beim logischen Schließen auftreten oder wenn mehrere Anforderungen gleichzeitig zu bewältigen sind (bspw. Baddeley, 1986). Zusätzliche Anforderungen, die beträchtlichen Koordinationsaufwand erfordern, behindern die Leistung beim konditionalen Schließen, beim syllogistischen Schließen sowie beim relationalen Schließen (für einen Überblick Knauff, 2006).

Wird nun abschließend der Blick auf kausales Schließen als Sonderform der Induktion gerichtet, so stellt sich die Frage, wie die Vorhersage von Konsequenzen beeinflusst wird, bzw. wie Kausalitäten erkannt werden (für einen detaillierten Überblick Waldmann, 2006; Buehner & Cheng, 2005; Waldmann & Pearl, 2000). Kausalität steht als unidirektionaler Zusammenhang im Gegensatz zur Korrelation. Waldmann und Holyoak (1992) gehen davon aus, dass wir durchaus sensitiv gegenüber der Unidirektionalität von Kausalitäten sind, das heißt, dass Ursachen Effekte beeinflussen können, Effekte aber nicht die Ursachen. Sie nehmen an, dass unser Vorwissen über kausale Zusammenhänge für die Annahme von Kausalitäten ausschlaggebend ist (Waldmann & Hagmayer, 2001). Dabei konnten folgende Einflussgrößen identifiziert werden (Hagmayer & Waldmann 2006; Waldmann & Holyoak, 1992; Thompson, 1994, 2000): Mit steigender Anzahl der den Effekt verhandelnden Faktoren nahm die Zustimmung zu Modus Ponens und Modus Tollens ab. Die Anzahl alternativer Ursachen hatte keinen Einfluss auf die Annahme oder Ablehnung beim Modus Ponens und nur einen kleinen Einfluss bei Modus Tollens. Entgegengesetzte Befunde ergaben sich für die Negation des Antecedens und die Affirmation der Konsequenz. Hier war die Zahl der bekannten weiteren Ursachen des Effektes entscheidend.

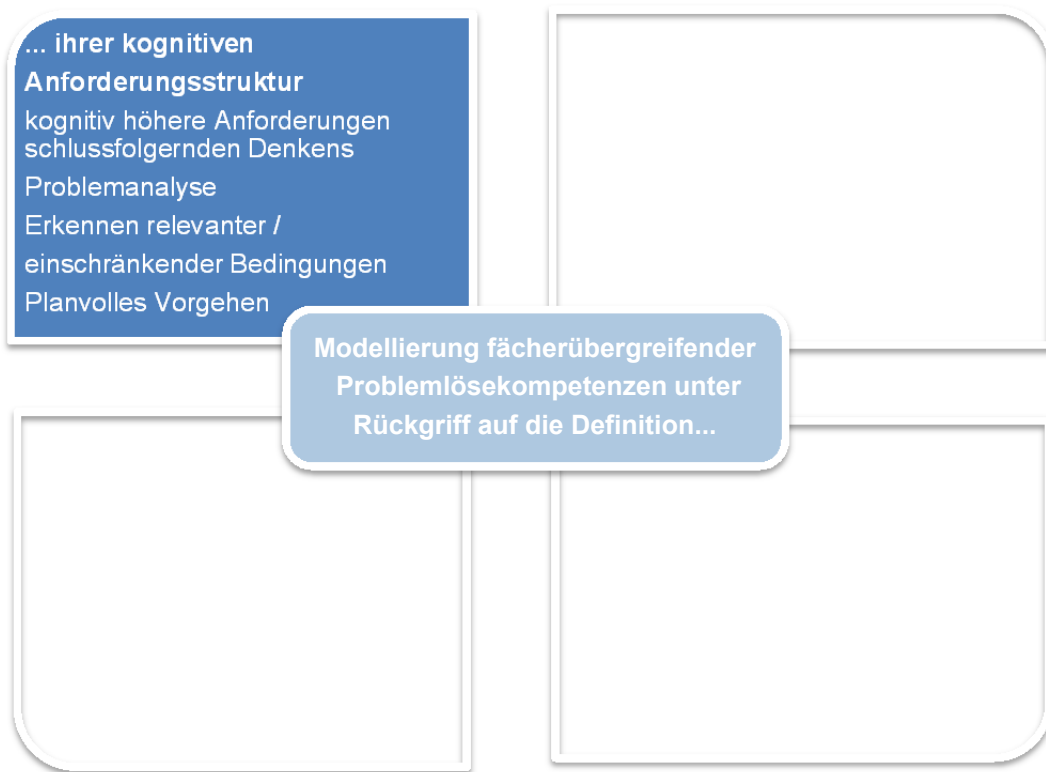
Weiterhin war die Vertrautheit des in der Aufgabe dargestellten Sachverhalts ausschlaggebend. Je besser ein Effekt bekannt ist, desto mehr verschiedene Ursachen können erinnert und desto mehr Fälle, in denen der Effekt trotz Ursache nicht aufgetreten ist, können aus dem Gedächtnis abgerufen werden. Dies führt auch dazu, dass mit hoher Vertrautheit eines Zusammenhangs die Notwendigkeit der Ursache für den Effekt abnimmt. Daher werden kausale Schlüsse bei bekannten Sachverhalten eher abgelehnt, während sie bei unbekanntem Sachverhalten eher angenommen werden (Cummins, 1995). Diese Befunde beziehen sich auf Sachverhalte, die Zusammenhänge nach dem Muster, „Wenn Ursache, dann Effekt“ präsentieren. Sachverhalte können allerdings auch in der Form vorliegen, „Wenn Effekt, dann Ursache“. Cummins geht hier nach der Theorie logischen Schließens davon aus, dass die Form der vorliegenden Prämissen für die Annahme der Kausalität ausschlaggebend ist (Cummins, 1995). Dies konnte allerdings nur teilweise bestätigt werden (Quinn & Markovits, 1998, Markovits & Potvin, 2001 nach Waldmann, 2006). Vielmehr deuten die Befunde darauf hin, dass die Stärke der Assoziation zwischen verschiedenen alternativen Ursachen und dem Effekt ausschlaggebend ist. Je stärker andere Faktoren mit dem Effekt assoziiert sind, desto geringer die Akzeptanz der Schlussfolgerung (Waldmann, 2006).

Zusammenfassung

Die Befunde zeigen, dass die Qualität des Problemlösens durch das Verständnis der jeweils gegebenen Problemsituation sowie die kognitiven Anforderungen des schlussfolgernden Denkens beeinflusst wird, bevor sie sich in der Angemessenheit der erreichten Lösung zeigt (auch Klieme et. al., 2001d). Anforderungen des logischen Schließens, das heißt, deduktives, induktives sowie kausales Schließen stellen dabei wesentliche kognitive Fertigkeiten im Rahmen des Problemlösens dar. Somit bestehen basale kognitive Fertigkeiten der Problemlösung darin, vorliegende Informationen abzugleichen und vorhandene Beziehungen zwischen Aussagen zu erkennen, um schließlich mit Blick auf eine zu treffende Schlussfolgerung zwischen vorhandenen Alternativen abzuwägen, wobei jeweils vorliegende Einschränkungen berücksichtigt werden müssen.

Für das fächerübergreifende Problemlösen stellt sich im Weiteren nun die Frage, was unter dem Begriff des Fächerübergreifenden zu verstehen ist. Wie fächer(un)spezifisch sind kognitive Fertigkeiten des logischen Schließens? Der Begriff des Fächerübergreifenden legt nahe, es gäbe ein Set fachunabhängiger Problemlösekompetenzen. Hierbei muss jedoch berücksichtigt werden, dass kognitive Operationen nicht dekontextualisiert, das heißt unabhängig vom Kontext gesehen werden können (Robbins & Aydede, 2009; Kirsh, 2009; Carey & Spelke, 2008; Graesser, Gholson & Houston, 1996; Chen & Holyoak, 1995; Schraw, Dunkle & Bendixen, 1995; Perkins & Simon, 1987, 1988). Um möglichst genau Problemlösefertigkeiten vorhersagen und beschreiben zu können, muss deshalb im Folgenden geklärt werden, in welcher Relation kognitive Fertigkeiten des Problemlösens zum Kontext, in dem sie erlernt werden sowie zum Kontext, in dem sie angewendet werden, stehen.

Abb. 2: Modell zur Beschreibung fächerübergreifender Problemlösekompetenzen



5.3 Kontextgebundenheit fächerübergreifender Problemlösekompetenzen

Für das fächerübergreifende Problemlösen (bzw. Schlüsselkompetenzen im Allgemeinen) ist die Frage nach deren Generalität, bzw. Spezifität relevant. Im Zentrum der Schlüsselkompetenzen steht, dass erst die übergreifende Anwendbarkeit den Schlüsselkompetenzen ihre „Schlüsselfunktion“ zuweist. Die Frage nach dem Kompetenztransfer, der Kontextgebundenheit, bzw. Generalität kognitiver Problemlösekompetenzen stellt damit eine der wichtigsten, bislang jedoch auch eine der umstrittensten Fragen dar (Nickerson, 1994; Perkins & Salomon, 1987, 1989; Nickers, Perkins & Smith, 1985):

„In fact, there is little agreement [...] about the nature of transfer, the extent to which it occurs, and the nature of its underlying mechanisms.“ (Barnett & Ceci, 2002, S. 612)

Dabei ist diese Frage nicht nur allgemeinspsychologisch interessant, sondern berührt auch anwendungsorientierte Aspekte, etwa, worauf überfachliche Expertise basiert und wie sie bestmöglich gefördert werden kann. Im Folgenden werden die wichtigsten Befunde zur fächer(un)abhängigen Verortung von Problemlösekompetenzen dargestellt. Die Annahme, dass die Effektivität fächerübergreifenden Problemlösens auf entweder domänenspezifischen, kontextgebundenen Fähigkeiten basiert oder aber auf generalisierbaren Kompetenzen, die in einem Kontext erworben und auf den anderen übertragen werden, entspricht jedoch eher Wunschenken als empirischer Evidenz

(Pressley et. al., 1987). Die Frage, welche Anteile fächerübergreifender Problemlösekompetenzen domänen(un)gebunden sind, muss vielmehr über das Zusammenspiel zwischen Transfer, Expertiseentwicklung und domänenspezifisch erworbenen Problemlösekompetenzen beantwortet werden.

Fächerübergreifendes Problemlösen unter dem Gesichtspunkt analogen Problemlösens und Transfers
Problemlösen ist als eine Suche nach Strukturen und Beziehungen zwischen den jeweils vorliegenden Problemelementen zu definieren. Beim fächerübergreifenden Problemlösen spielt nun der Aspekt des Neuen eine entscheidende Rolle. Im Mittelpunkt steht dementsprechend die Fähigkeit, neue Probleme unter Anwendung domänenspezifisch erworbener kognitiver Problemlösekompetenzen lösen zu können. Damit werden beim fächerübergreifenden Problemlösen bereits vorhandene kognitive *skills*, kognitive Schemata, Strukturen und Beziehungen als Ausgangspunkte dafür herangezogen, um in unbekanntem Problemsituationen bekannte Strukturen und Beziehungsmuster zu erkennen (Nokes, 2009; Holyoak, 2005; Bassok, 2003; Mayer & Wittrock, 1996; Van Leer, 1987). Fächerübergreifende Problemlösekompetenzen könnten damit auch als analoge Problemlöse- bzw. Transferkompetenzen angesehen werden. Ins Zentrum rückt damit die Frage, unter welchen Umständen Effekte von Vorwissen auf neue Problemsituationen festgestellt werden können (Mayer, 2010; Mayer & Wittrock, 2006). Für einen detaillierteren Blick wird auf Nokes (2009), Holyoak (2005), Bassock (2003), Barnett und Ceci (2002), Bransford, Brown und Cocking (2000), Novick (1995), Mayer und Wittrock (1996) oder auch Salomon und Perkins (1989) verwiesen.

Die Nutzung bekannter Episoden des Problemlösens bei der Bearbeitung neuer Aufgaben im Sinne analogen Transfers hängt von der erfolgreichen Bewältigung verschiedener komplexer Teilprozesse ab, auf die in aller Ausführlichkeit an dieser Stelle nicht eingegangen werden kann (Repräsentation des Zielproblems, Aktivierung eines Quellproblems, Auswahl des Quellproblems, Mapping und Inferenz, Mustererkennung und Adaption; für eine detaillierte Zusammenfassung empirischer Evidenzen vgl. Scheiter, 2003). Wird neues Problemlösen durch Vorangegangenes erleichtert, spricht man von positivem Transfer (Mähler & Stern, 2006). Wirkt sich das frühere Lernen hemmend auf die Lösung neuer Probleme aus, hat negativer Transfer stattgefunden (ebd.). Negativer Transfer bedeutet, dass die Veränderung bereits verfügbarer Kenntnisse und Fertigkeiten die Lösung neuer Probleme beeinträchtigt. So ist beispielsweise häufig nach der Vermittlung von Regeln bzw. Problemlösestrategien zu beobachten, dass es anschließend zu Übergeneralisierungen und damit Fehllösungen neuer Aufgabenanforderungen kommt (Mähler & Stern, 2006).

Der Transfer von Problemlösekompetenzen auf unbekannte Problemstellungen basiert grundsätzlich auf dem Ausmaß der Ähnlichkeit der unbekanntem Situation mit bekannten Situationen, wobei relevante Bedingungen positiven Transfers die Ausprägung und Wahrnehmung von

Strukturmerkmalen bzw. strukturellen Ähnlichkeiten, die Kombination von Strukturmerkmalen mit Oberflächenmerkmalen und die Verfügbarkeit domänenspezifischen Vorwissens sind (vgl. Kap. log. Schließen). Die Übertragung bekannter Reaktion auf eine Problemsituation in Abhängigkeit ihrer Ähnlichkeit mit unbekanntem Problemsituationen findet deshalb umso eher statt, je ähnlicher die Situationen sind (Greeno, Collins & Resnick, 1996). Analogien sind dann vorhanden, wenn zwei Situationen gleiche oder ähnliche Muster in ihren konstituierenden Elementen besitzen (Holyoak, 2005; Bassok, 2003).

Wie läuft analoges Problemlösen ab? Das vorliegende neue Problem, die Zielanalogie (target) stellt den Ansatzpunkt dar, um nach bereits vorhandenen, domänenspezifisch erworbenen Basisanalogien (source) zu suchen. Die Basisanalogie wird herangezogen, um Rückschlüsse über das vorliegende neue Problem zu ziehen. Ausschlaggebend für das Finden einer Basisanalogie sind bestimmte Schlüsselähnlichkeiten. Diese werden jedoch, wie Genter (1982) zeigen konnte weniger in den einzelnen Problemelementen gesucht werden, sondern vielmehr in den von ihr angesprochenen Domänen.

Welchen Einfluss haben Konzeption und Struktur von Ziel- und Basisanalogie auf den Erfolg der Transferleistung? Hier zeigen zahlreiche Befunde, dass analoges Problemlösen und die entsprechende Transferleistung maßgeblich von Oberflächen- und Struktur(un)ähnlichkeiten bei *Source*- und *Target*-Analogien abhängen (Nokes, 2009; Holyoak, & Koh, 1987; Spencer & Weisberg, 1986; Gick & Holyoak, 1983; Genter, 1982, 1983; Reed et al. 1974). Analoges Problemlösen wird wesentlich von Oberflächenmerkmalen wie dem Problemkontext, den jeweiligen Textmerkmalen oder auch dem Probleminhalt beeinflusst (Bassok, 2003; für eine Übersicht Reeves & Weisberg, 1994).

Dabei interferieren vor allem gleiche, bzw. ähnliche (markante) Oberflächenmerkmale von Ziel- und Basisanalogie mit Transfer, da sie den Blick auf zugrundeliegende strukturelle (Un)Ähnlichkeiten zwischen Problemanalogien verdecken. Hilfestellungen und Hinweise auf Transfer können hier zwar hilfreich sein und Transfer ermöglichen (Gick & Holyoak, 1983; auch Spencer & Weisberg, 1986). Spontaner positiver oder auch negativer Transfer findet dennoch umso eher statt, je ähnlicher Oberflächenmerkmale in Ziel- und Basisanalogie sind. Die Struktur der Analogien nimmt insofern Einfluss, dass Transfer eher verhindert wird, wenn strukturelle Unähnlichkeiten vorhanden sind: In diesem Fall ist selbst ein Hinweis auf eine Transfermöglichkeit weniger wirksam (Holyoak & Koh, 1987; Reed, Ernst & Banerji, 1974). Interessant ist im Hinblick auf Problemlösen deshalb vor allem die Manipulation der Oberflächenstruktur zusätzlich zur Lösungsstruktur. Dabei ist zu beobachten, dass die von Novizen präferierte Oberfläche allmählich während des Lernens in komplexere Strukturen übergeht (Brunstein, 2001).

Transfer kann dementsprechend als Funktion von Analogieelementen beschrieben werden. In Bezug auf die Frage der Förderung von Transfer konnte zwar generell ein Einfluss von Vorwissen auf die Transferleistung festgestellt werden (Gick & Holyoak, 1980), allerdings in Abhängigkeit der

empfundene Nähe zwischen Ziel- und Basisanalogie sowie der gegebenen Hilfestellung. Die Nähe zwischen Ziel- und Basisanalogie spielt zudem auch dann eine Rolle, je mechanischer die eingeübten Problemlösestrategien sind (Mayer, 1981). Genter (1983) betont deshalb im Zusammenhang mit analogem Problemlösen, dass die Transferleistung im Wesentlichen davon abhängt, ob der mapping-Prozess zwischen Ziel- und Quellanalogie erfolgreich abläuft (Bassok, 2003; Barnett & Ceci, 2002; Catrambone & Holyoak, 1989, nach Barnett & Ceci, 2002; Anderson, 1996; Niedelman, 1991).

Die Vertrautheit des Kontextes spielt dementsprechend eine große Rolle (Brown et. al., 1993; Schliemann und Magalhães, 1990; Schliemann & Nunes, 1990 nach Barnett & Ceci, 2002). Relevant wird dabei weiterhin, ob eine Ankeranalogie vom Problemlöser vor allem als brauchbar und plausibel angesehen wird (Brown & Clement, 1989). Bransford, Brown und Cocking (2002) verweisen darauf, dass die Förderung der Transferleistung deshalb vor allem damit beginnt, den Prozess des Erkennens von Ankeranalogien anzuleiten und auf ein Verständnis des Prozesses des analogen Problemlösens und des Transfers hinzuwirken. Dabei ist die Vermittlung von Transferkompetenzen anhand eines umfassenden Aufgabenspektrums besser als anhand beispielhafter Einzelaufgaben (Lim, Reiser & Olina, 2009).

Wie der Überblick zeigt, ist unumstritten, dass Transfer von verschiedenen Bedingungen abhängt. Allerdings legt die große Anzahl an unterschiedlichen Befunden sowie vor allem der Nachweis negativen oder nicht gelungenen Transfers nahe, dass Transferbedingungen komplex sind (bspw. LOGO, Bransford et. al., 2002; Luchins, 1942; für einen Überblick Anderson & Singley, 1989). Im Abgleich mit dem Vorliegen gelungenen Transfers (u.a. Bearman, Ball & Ormerod, 2007; Stark, 2000; Schunn & Anderson, 1999; Schwartz, & Goldman, 1996; Fong & Nisbett, 1991; Lehmann & Nisbett, 1990; Lehmann, Lempert & Nisbett, 1988) wurden deshalb Transfermodelle (bspw. ADAPT, Jelsma, Marrienboer & Bijlstra, 1990) sowie Transfertypologien und -taxonomien (bspw. Gagné, 1977; auch Langley, 1985, Brooks & Dansereau, 1987, Genter & Toupin, 1986; Perkins & Salomon, 1987) erarbeitet, um transferrelevante Bedingungen und Kompetenzen zu systematisieren.

Systematisierungsansätze

Ein Blick auf die Kategorisierungsansätze zeigt, dass diese von der Distanz zwischen Basis- und Zielproblem ausgehen und sich an den bereits genannten Ähnlichkeitsmaßen orientieren. Sie unterscheiden zwischen *generellem* und *spezifischem Transfer*, lassen jedoch auch *Mischmodelle* zu und beleuchten die Frage, inwieweit Ähnlichkeiten sowohl von Kompetenzen als auch Problemstellungen zu Transfer führen:

So geht die Annahme *generellen Transfers* davon aus, dass eine Disziplin zwar die Kompetenz A vermittelt, Kompetenz A jedoch auch zugleich Kompetenz B fördert. Mit generellem Transfer wurde beispielsweise oftmals für die Relevanz des Erlernens von Latein als grundlegender Kompetenz

argumentiert. Thorndike und Woodworth widerlegten dies allerdings bereits 1901: Es zeigte sich, dass Studierende, die Latein erlernten, nicht besser in Buchhaltung waren als Studierende ohne Vorkenntnisse in Latein. Auch Anderson und Singley (1989) fanden wenige Evidenz für allgemeinen Transfer. Demgegenüber steht das Erklärungsmodell *spezifischen Transfers*, das propagiert, dass das Erlernen der Kompetenz A dabei hilft, Aufgabe B zu bearbeiten, da einige Elemente ähnlich oder gleich sind. Perkins und Salomon (1989) bezeichnen dies auch als *vertikalen Transfer*. Das Erlernen von Latein wäre demzufolge für das Erlernen romanischer Sprachen deshalb hilfreich, da einige Elemente (bspw. Wörter, Grammatik) ähnlich sind. Der *spezifische Transfer von allgemeinen Prinzipien* findet wiederum statt, wenn die Strategie für Aufgabe A auch in Aufgabe B verwendet werden kann, obwohl Aufgabe B keine ähnlichen Elemente teilt. Perkins und Salomon (1989) bezeichnen dies als *horizontalen Transfer*. Hier wird Transfer als generelles Prinzip auf Basis gemeinsamer Schemata, Muster und Beziehungen angesehen und weniger als spezifisches Verhalten. *Gemischte Modelle* gehen davon aus, dass spezifischer und genereller Transfer interagieren: So wird, wenn bspw. von Aufgabe A das Prinzip erkannt und abstrahiert wird, dieses Prinzip auf Aufgabe B angewendet. Spezifisch ist der Transfer deshalb, da die Aufgaben A und B über den ähnlichen oder gleichen Mechanismus bearbeitet werden, was jedoch übertragen wird, ist ein abstraktes Prinzip. Ein Beispiel wäre das Erlernen des Schreibens von Zusammenfassungen, die nach dem gleichen Prinzip in sehr verschiedenen Situationen angewendet werden.

Diese ebenso wie Modelle, die sich mit der Frage der Distanz zwischen Ziel- und Basisanalogien beschäftigen (bspw. Langley, 1985, Brooks & Dansereau, 1987, Genter & Toupin, 1986; Gagné, 1977) geben zwar Anhaltspunkte über die Ähnlichkeiten zwischen Kompetenzen bzw. Problemstellungen. Sie bieten allerdings wenig Anhaltspunkte wann, in welchen Situationen oder auch bei welchen Problemstellungen welche Art von Transfer geleistet werden muss, bzw. was transferiert wird. Diesbezüglich legen Barnett und Ceci (2002) ein Modell vor, anhand dessen Transfer in Bezug zu relevanten Transferbedingungen gesetzt wird.

Grafik 1: Transfertaxonomie in Abhängigkeit von Inhalt und Kontext

Content: What transferred?			
Learned skill	Procedure	Representation	Principle or heuristic
Performance concept	Speed	Accuracy	Approach
Memory demands	Execute only	Recognize and execute	Recall, recognize, and execute

Context: When and where transferred from and to		
	Near Transfer	Far Transfer
Knowledge domain	Mouse vs. rat	Science vs. art
Physical context	Same room at school	School vs. the beach
Temporal context	Same session	Years later
Functional context	Both clearly academic	Academic vs. play
Social context	Both individual	Individual vs. society
Modality	Both written same format	Lecture vs. wood carving

(Zusammengefasst nach Barnett & Ceci, 2002, S. 621)

Fächer(un)abhängigkeit fächerübergreifender Problemlösekompetenzen

Gibt es ein Set domänenunspezifischer Problemlösekompetenzen, die in domänenunabhängigen Problemlösesituationen virulent werden? Unterschiedliche Studien beleuchten diese Frage. Sowohl die Suche nach einem *General Problem Solver* der Informationsverarbeitungstheoretiker (*power-generalizy-trade-off*, Perkins & Salomon, 1989), als auch die Suche nach einer übergeordneten Rolle von domänenspezifischem Wissen oder allgemeinen Problemlöseheuristiken (Schunn & Anderson, 1999; Niedelman, 1991; Schoenfeld, 1985) zeigt, dass sowohl die Annahme kontextunabhängiger als auch kontextabhängiger Problemlösekompetenzen zu unbefriedigenden Befunden führt. Die Befunde zeigen, dass die Effektivität fächerübergreifenden Problemlösens weder auf domänenspezifischen, kontextgebundenen Fähigkeiten basiert, noch auf generalisierbaren Kompetenzen. Die Rolle des Transfers, bzw. analogen Problemlösens muss deshalb in jedem Fall mitberücksichtigt werden (Lehman, Lempert & Nisbett, 1988; Nickerson, et. al. 1985; Hayes & Simon, 1977).

Die Eingangsfrage ob fächerübergreifende Problemlösekompetenzen domänengebunden oder domänen(un)gebunden sind, muss dementsprechend über die Abhängigkeiten zwischen Transfer, Expertise und allgemeinen Problemlösekompetenzen und -heuristiken beleuchtet werden. Interessant ist deshalb die Frage, was passiert, wenn Experten mit unbekanntem Problemen konfrontiert werden. Wie kann das Zusammenspiel von domänenspezifischem Wissen, Transfer und Problemlösekompetenzen beschrieben werden (bspw. Connell, Sheridan & Gardner, 2003; Perkins & Salomon, 1989)? Dahingehend zeigt sich, dass Experten, konfrontiert mit atypischen Problemen, wieder auf allgemeinere Strategien zurückgreifen, intuitive Modelle oder Extrembeispiele benutzen oder die unbekanntem Probleme auf einfachere Probleme herunterbrechen (bspw. Clement, 1980). Auch Schunn und Anderson (1999) kommen zu ähnlichen Ergebnissen. Sie ziehen den Schluss, dass allgemeine Problemlösestrategien zwar über domänenspezifisches Wissen moderiert werden, neue Problemstellungen jedoch zu einem Rückgriff auf allgemeinere Strategien führen (vgl. auch SOAR, Laird, Rosenbloom & Newell, 1984; ACT-R, Anderson). Schoenfeld (1982) fragt deshalb danach, inwieweit die Vermittlung allgemeiner Problemlöseheuristiken zur Verbesserung domänenspezifischer Kompetenzen führt und kann zeigen, dass *weak methods*, die kontextualisiert dargeboten werden, durchaus eine Verbesserung von Mathematikleistungen zur Folge haben. Nokes (2009) kann nachweisen, dass abhängig von der jeweiligen Problemanforderung Transfer vorwissensgebunden abläuft. Vorwissen nimmt demzufolge zwar Einfluss, allerdings auch die Konzeption des Problems. Insgesamt lässt sich damit zusammenfassen, dass sich Problemlösen über das Anforderungsniveau und das Vorwissen entscheidet.

Tragen dementsprechend fächerübergreifende Problemlösekompetenzen lediglich einen verheißungsvollen Namen und sind letztlich doch nur eine Frage der Domäne oder des Problems? Die Antwort, die zwar die Komplexität des Sachverhaltes berücksichtigt, allerdings für die Praxis wenig hilfreich ist, lautet wohl: Das hängt davon ab. Ein ausschließlicher Fokus auf die Vermittlung kognitiver Kompetenzen und Problemlöseheuristiken negiert die Rolle der domänengebundenen Expertise. Expertise wird ihrerseits unterstützt durch den Rückgriff auf allgemeine Problemlösekompetenzen. Zudem müssen die Bedingungen des Transfers domänenspezifisch erworbener, jedoch fächerübergreifend relevant werdender Problemlösekompetenzen Berücksichtigung finden (u.a. Barnett & Ceci, 2002). Fächerübergreifende Problemlösekompetenzen agieren demzufolge weder unabhängig noch abhängig von domänenspezifischen Kompetenzen, sie werden auch nicht unabhängig von der Disziplin bei jedem Problem gleich wirksam:

„You don't pick up a baby [...] as a basket of laundry.“ (Perkins & Salomon, 1989, S. 23)

Damit wird die Frage nach den Anforderungen, die das Problem stellt, zentral.

5.4 Modellierung fächerübergreifenden Problemlösens in Abhängigkeit des Problems

Die Definition fächerübergreifender Probleme als Probleme realistischer, ganzheitlicher Natur (Baumert, Klieme, Neubrand & Lüdtke, 2001) zeigt, dass die Probleme, die darunter verstanden werden, zwar logisches Schlussfolgern erfordern, jedoch andere Anforderungen stellen als die einfach strukturierten Syllogismen. Daraus leitet sich die Frage ab, wie Probleme unterschieden werden können und wie das Problem Einfluss auf seine Lösung nimmt.

Wie nimmt das Problem Einfluss auf seine Lösung?

„[W]hat makes some problems really hard“ (Kotovsky & Simon, 1990)? Eine Vielzahl an Studien beschäftigte sich mit der Prognose möglicher Schwierigkeiten beim Problemlösen (Mayer, 2010; Pretz, Naples & Sternberg, 2004; Funke, 2003; Jonassen, 2000; Kluwe, 1995; Schraw, Dunkle & Bendixen, 1995; Greeno & Simon, 1988; Kitchener, 1983; für fächerübergreifende Probleme Baumert et al., o. J.; Fleischer et. al., 2010). Dabei zeigt sich, dass die Unterscheidung zwischen gut und schlecht definierten Problemen eine grundlegende Differenzierung und damit eine Prognose möglicher Schwierigkeiten erlaubt. Die Charakteristika, anhand derer sich das Problem in gut oder schlecht definiert unterscheiden lässt, sind,

- ob der Problemausgangspunkt definiert ist,
- das Ziel, das es zu erreichen gilt, klar ist,
- die Informationen zur Beantwortung der Fragestellung und damit zur Zielerreichung vorhanden sind und
- ob die Hürden, die der Zielerreichung entgegenstehen, ebenfalls klar sind.

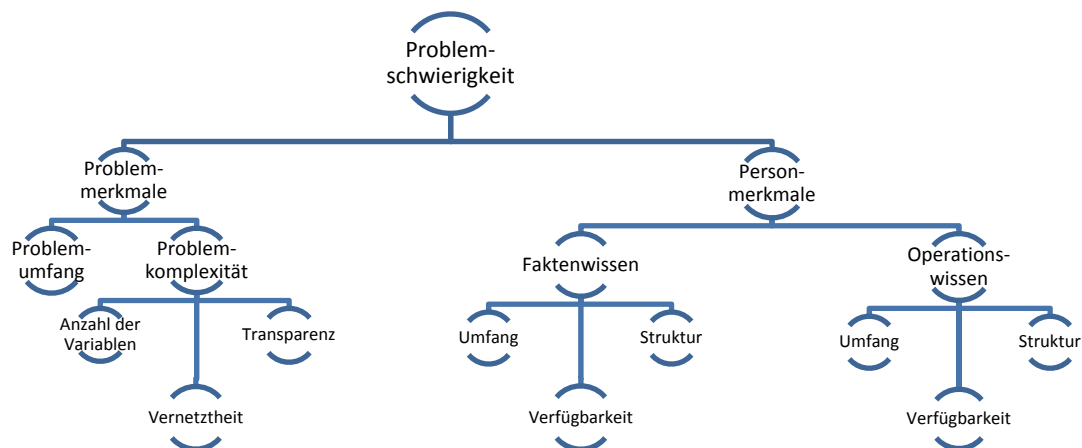
Ein Vergleich gut und schlecht definierter Probleme zeigt, dass gut definierte Probleme besser über fachunspezifische und allgemeiner anwendbare Problemlösekompetenzen gelöst werden können als schlechter definierte Probleme, die domänenspezifischere Kompetenzen verlangen und als eher wissensabhängig gesehen werden (Greeno & Simon, 1988). Darüber hinaus erfordern schlecht definierte Probleme die Anwendung mehrerer kognitiver Operationen (Kluwe, 1995), die Arbeitsgedächtnisbeanspruchung steigt, die Anforderung mehr Variablen zu manipulieren, zu behalten und zu bearbeiten beansprucht mehr Kapazität (Schraw, Dunkle & Bendixen, 1995). Dabei nimmt vor allem die Kontextualisierung Einfluss: Die zur Lösung von gut definierten Problemen notwendigen Fertigkeiten sind kontextbasierter als diejenigen, die zur Lösung von schlecht definierten Problem notwendig sind (auch Karmiloff-Smith, 1992; Kitchener, 1983). Schlecht definierte Probleme erfordern vor allem Transferkompetenzen (auch Chandler et. al. 1990; Chi & Glaser, 1988). Darüber hinaus wird bei schlecht definierten Problemen auch der Einfluss von *beliefs* relevanter (Jonassen, 2000; Dörner, 1989), da sie die Mitberücksichtigung der epistemischen Struktur der Probleme oder auch der Legitimität der kognitiven Prozesse, die zur Lösung eingesetzt werden erfordern. Die Schwierigkeit eines Problems wird damit in erster Linie über seine Komplexität beeinflusst, das

Arbeitsgedächtnis wird als Invariante des Problemlösens gesehen. Über welche weiteren Problemkomponenten kann das Problemlöseverhalten und der Problemlöseerfolg vorhergesagt (Ericsson & Hastie, 1994; Kotovsky & Simon, 1990), bzw. die Schwierigkeit eines Problems variiert werden?

Fächerübergreifendes Problemlösen: Eine Frage des Problems?

Es liegen zahlreiche Klassifikationsansätze zur Charakterisierung von Problemen vor: So wird nach Problemtypus, Problemcharakteristik und Art der verlangten Informationsverarbeitung unterschieden (Arlin, 1989), nach der Art der Barriere, Bekanntheit der Mittel sowie Klarheit der Zielkriterien Dörner (1976; 1989) oder auch nach der Lücke im Problemraum (Lür & Spada 1990; nach Funke, 2003). Funke (1990) oder auch Hussy (1984) widmen sich komplexen Problemen und ihrer Differenzierung. Die nachstehende Abbildung fasst in einem Überblick relevante Einflussfaktoren auf die Problemschwierigkeit zusammen.

Grafik 2: Determinanten der Problemschwierigkeit



(nach Hussy, 1984)

Im Folgenden soll der komplexeren Kategorisierung von Jonassen (2000) gefolgt werden. Sie versucht eine Integration der vorhandenen Modelle, bleibt jedoch bei pragmatisch-funktionalen Kategorien und kann so einerseits die Vielfalt an Problemen differenziert abbilden, andererseits dennoch anwendbar bleiben.

Problemlösen definiert Jonassen (2000) als abhängig von folgenden Komponenten:

Problemtypus	<ul style="list-style-type: none"> • Definiertheit, bzw. Strukturiertheit des Problems • Komplexität
--------------	--

	<ul style="list-style-type: none"> • Domänenspezifität
Repräsentationsform	<ul style="list-style-type: none"> • Kontext des Problems, hier verstanden als sozialer, historischer oder auch kultureller Kontext • Hinweisreizen (<i>cues</i>) • Problemmodalität
Individuelle Unterschiede	<ul style="list-style-type: none"> • Expertise • strukturelles, prozedurales und konzeptuelles Wissen • <i>Reasoning</i> • kognitiver Stil • Selbstwirksamkeit und Motivation

Dass Probleme mit Blick auf ihre *Definiertheit* variieren, wurde bereits vorgestellt: Gut strukturierte Probleme beinhalten alle zur Lösung relevanten Problemelemente, erfordern die Anwendung einiger weniger *weak-methods* und haben eine Lösung, die sowohl bekannt sein kann, als auch verständlich ist und bei der die Beziehung zwischen Problemzuständen bekannt oder wahrscheinlich ist. Weniger gut, bzw. schlecht strukturierte Probleme zeichnen sich dadurch aus, dass weniger, nicht alle oder auch keine Problemelemente bekannt sind, durchaus auch verschiedene Lösungen und Lösungswege möglich sind, oder bspw. auch gar keine Lösung vorhanden ist. Ein Beispiel hierfür sind moralische Problemstellungen. Hier zeichnet sich auch ein weiteres Kriterium ab: Weniger gut, bzw. schlecht strukturierte Probleme erlauben oftmals verschiedene Kriterien zur Evaluation des Lösungserfolgs, eine zusätzliche Unsicherheit besteht dementsprechend auch dahingehend, entscheiden zu müssen, welche der verschiedenen Strategien und Operatoren angewendet werden sollen, um zu einer zufriedenstellenden Problemlösung zu gelangen (bspw. Sahelzonen-Szenario, Dörner, 1989). Schlechter strukturierte Probleme implizieren damit oftmals auch, Entscheidungen auf Basis persönlicher Dispositionen und Einstellungen zu fällen. Jonassen zählt zum Aspekt der Definiertheit/Strukturiertheit auch die Verfügbarkeit von Informationen sowie die kontextuelle Vielfalt. Hier zeigt sich, dass die Charakteristika nicht unabhängig voneinander sind, da die kontextuelle Vielfalt und Verfügbarkeit von Informationen auch die Komplexität eines Problems beeinflusst.

Die *Komplexität* der Probleme definiert sich vor allem über die Anzahl der Problemelemente, bzw. Problemvariablen sowie den Grad ihrer Verknüpftheit. Zudem wird ein Problem als leichter oder schwieriger angesehen, je nachdem welche funktionale Beziehung zwischen den Problemvariablen herrscht und wie stabil deren Eigenschaften sind. Vielfältige, dynamische und intransparente Zusammenhänge zwischen Problemkomponenten erschweren die Abschätzung, inwieweit sich Veränderungen einstellen, wenn bestimmte Problemvariablen verändert werden. Das heißt, das

Kriterium der Komplexität eines Problems beschreibt, wie viele und wie klar Problemkomponenten implizit oder explizit dargestellt sind. Der Aspekt der Komplexität stellt den entscheidenden Faktor für die Schwierigkeit des Problems dar. Intuitiv variieren auch Versuchspersonen vor allem die Vernetztheit der Problemkomponenten, um den Schwierigkeitsgrad von Problemen zu erhöhen (Nährer, 1988).

Letztlich lassen sich Probleme auch anhand ihrer *Domänenspezifität* und ihrer *Repräsentationsform* unterscheiden. Den Aspekt der *Domänenspezifität* ergänzt Jonassen über die dichotomen Kriterien Abstraktheit bzw. Situiertheit des Problems.

Unter der *Repräsentationsform* eines Problems ist nicht nur die Präsentationsform der Probleme an sich gemeint (visuelle, verbale, numerische, symbolische, grafische Darstellungsmodalität). Vielmehr ist auch die kognitive Repräsentation gemeint.

Daneben liegen Kriterien zur Unterscheidung zwischen fachspezifischen und fächerübergreifenden Problemen vor (Fleischer, Wirth, Rumann & Leutner, 2010):

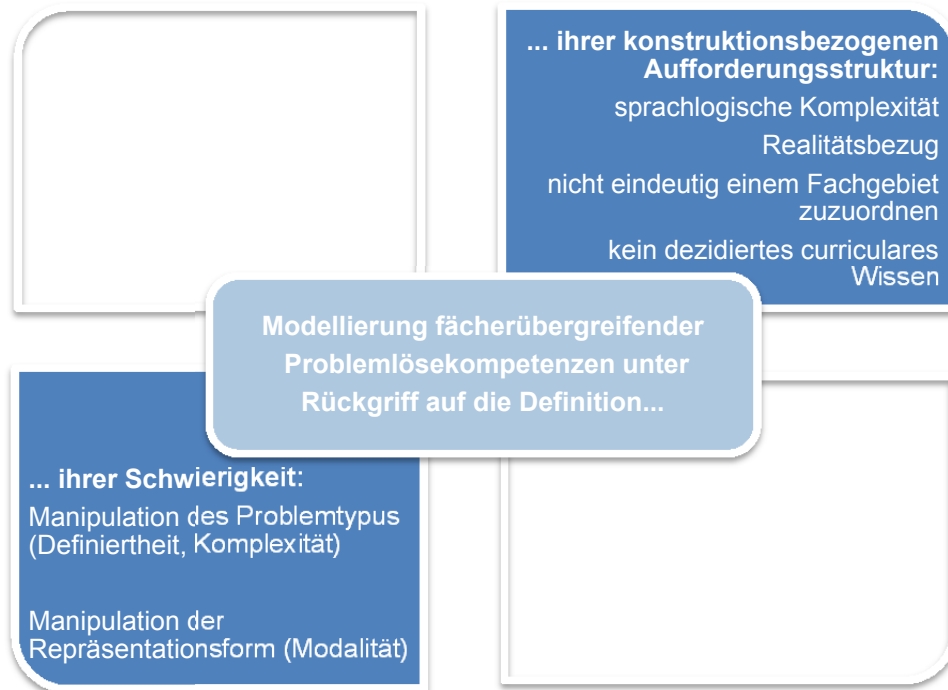
Fächerübergreifende Problemlöseaufgaben sind demnach eher als *Aufforderungen* denn als Fragestellungen formuliert und lassen sich dementsprechend auch *kognitiv höheren Anforderungen* zuordnen als fachgebundene Aufgaben (bspw. Mathematik). Fächerübergreifende Problemlöseaufgaben weisen eine signifikant höhere *sprachlogische Komplexität* auf, wohingegen fachgebundene Aufgaben einen höheren Formalisierungsgrad beinhalten (fachsprachliche Ausdrücke). Fächerübergreifende Problemlöseaufgaben sind *realistischer* formuliert und weisen eine signifikant höheren Bezug zu Familie, Freizeit, Beruf, Schule, Umwelt auf als fachspezifische Aufgaben. Fächerübergreifende Problemlöseaufgaben lassen sich gegenüber fachspezifischen Aufgaben signifikant *weniger eindeutig einem Fachgebiet zuordnen*. In Bezug auf die Funktion des Aufgabenkontextes sowie Aspekte wie Motivation oder auch Handlungsanleitung wurde kein bedeutsamer Unterschied festgestellt.

Fächerübergreifende Problemlöseaufgaben erfordern dementsprechend *weniger dezidiertes curriculares Wissen* als fachspezifische Aufgaben, die *kognitive Komplexität*, der Anspruch an planvolles, strategisches Vorgehen, wird höher eingestuft als bei fachgebundenen Aufgaben. Fächerübergreifende Probleme erfordern bedeutsam mehr *Handlungs- und Planungswissen* als fachgebundene, das Sachwissen ist auf einer allgemeineren Ebene angesiedelt. Die zur Lösung notwendigen Planungsschritte sind bei fächerübergreifenden Problemstellungen komplexer, es müssen mehr relevante und einschränkende Bedingungen erkannt werden. Ob fächerübergreifende Probleme mehr Anforderungen an das Verstehen der Situation stellen als fachgebundene konnte nicht gezeigt werden.

Für die vorliegende Arbeit wurden ausgehend von diesen Befunden die nachstehenden Ansatzpunkte zur Manipulation der Schwierigkeit fächerübergreifender Probleme herausgegriffen. Sie sollen eine

möglichst genaue Beschreibung und Prognose fächerübergreifender Problemlösekompetenzen erlauben:

Abb. 3: Modell zur Beschreibung fächerübergreifender Problemlösekompetenzen



5.5 Fächerübergreifendes Problemlösen: Eine differentialpsychologische Sicht auf das Konstrukt

Einfluss auf die Schwierigkeit von Problemen nehmen auch Bedingungen, die in der Person des Problemlösers liegen (Jonassen, 2000; Hussy, 1984). Wie komplex die Einflüsse von Lernercharakteristika auf Leistung im Allgemeinen sind, stellten u.a Wang, Haertel und Walberg (1990, 1993; auch Souvignier & Gold, 2006) dar. Im Folgenden werden zentrale individuelle Einflussfaktoren auf fächerübergreifendes Problemlösen dargestellt, wobei weniger einem eigenschaftsorientierten klassifikatorischen Ansatz (bspw. Problemlösetypen, kognitive Stile, habituelle Neigungen beim komplexen Problemlösen; Lantermann, Papen & Siebler, o. J.) gefolgt wird, sondern unter Berücksichtigung einer interaktionistischen Sichtweise individuelle Faktoren und ihr Einfluss auf Problemlösen vorgestellt werden. Damit soll der Blick auch auf nicht-kognitive Ansatzpunkte zur Förderung von fächerübergreifendem Problemlösen gerichtet werden.

Im Folgenden wird der Zusammenhang von Intelligenz und Problemlösen dargestellt und auf die Problematik der Abgrenzung beider Konstrukte Bezug genommen (u.a. Sternberg, 2005, 1982; Frensch & Funke, 1995). Weiterhin wird der Einfluss von Expertise und Vorwissen nochmals aufgegriffen (u.a. André & Fernand, 2008; Chi, 2006; Novick & Bassok, 2005; Ericsson, 2003b;

Sternberg & Frensch, 1992; Chi, Glaser & Rees, 1982) und die Relevanz epistemologischer Überzeugungen (Hofer, 2008, 2001; auch Heigl & Thomas 2013; Bromme, Pieschl & Stahl, 2010; Buehl, 2008; Muis, Bendixen & Haerle, 2006; Schommer-Aikins, Duell & Hutter, 2005; Bendixen & Rule, 2004; Sandoval & Morisson, 2003; Sternberg, 1994) sowie der motivationalen Komponenten der Selbstwirksamkeit und des Studieninteresses (u.a. Caprara, Vecchione, Alessandri, Gerbiono & Barbaranelli, 2011; Mayer, 2010, 1998; Schwarz & Skurnik, 2003; Pintrich & De Groot, 1990) im Zusammenhang mit Problemlösen dargestellt. Für einen Überblick über weitere personbezogene Einflussfaktoren wird auf Poropat (2011), Funke (2003), Mayer (2003), Schwarz und Skurnik (2003) verwiesen.

Fächerübergreifendes Problemlösen und Intelligenz

Zu Beginn wird Bezug genommen auf eine der wichtigsten Kontroversen in der Forschungsgeschichte des Problemlösens: Problemlösen und Intelligenz. Typische Aufgabenstellungen zur Erfassung von (Test)Intelligenz als logisches Schlussfolgern oder *reasoning* umfassen beispielsweise das Erkennen von Gesetzmäßigkeiten in numerischen oder symbolischen Darstellungen oder das Finden von Analogien. Typische Aufgabenstellungen zum Problemlösen liegen hingegen nicht vor – bzw. in Form klassischer Probleme wie den bereits berichteten. Problemstellungen variieren von gut strukturierten bis hin zu komplexen Szenarien (bspw. Moro/Lohausen, Dörner, 1989, 1976). Dieses Beispiel verdeutlicht, dass neben gemeinsamen Komponenten beim Problemlösen weitere Komponenten, insbesondere Regulationsfähigkeiten oder auch domänenspezifisches Wissen relevant werden. Die Frage ist, inwieweit Gemeinsamkeiten und Unterschiede zwischen Intelligenz und Problemlösen zu finden sind. Die Problematik der Differenzierung zeigt sich dabei bereits bei der Definition der Konstrukte: „Indeed, when two dozen prominent theorists were recently asked to define intelligence, they gave two dozen somewhat different definitions.” (Sternberg & Detterman, 1986, nach Neisser, 1996 S. 77). Dabei definieren letztlich die formalen Komponenten eines Problems wie eng der Zusammenhang mit Intelligenz ist (Sternberg, 2005) – für eine enge Beziehung der Konstrukte sorgt der Zusammenhang über den *reasoning* Faktor.

Bis heute ist umstritten, wie hoch die beiden Konstrukte korrelieren (Frensch & Funke, 1995; Baumert et. al., n.d.). Dabei zeigen sich höhere Korrelationen bei analytischem Problemlösen und geringere Korrelationen bei komplexem Problemlösen (Funke, 2003). Sternberg (1997) versucht die Debatte insofern zu klären, dass er Intelligenz – unter Berücksichtigung einer Differenzierung zwischen Testintelligenz und intelligentem Verhalten – zwar mit Problemlösen gleichsetzt, dabei jedoch verschiedene Teilkomponenten unterscheidet und dem Konstrukt Problemlösen eine praktische Komponente zuspricht (Sternberg, 1997, 1999, 2003; Sternberg & Kaufmann 1998). Inwieweit können Wirkungszusammenhänge zwischen Intelligenz und Problemlösen beschrieben werden?

Als eine der ersten Arbeiten zum Thema beschäftigt sich Putz-Osterlohs (1981) Trainingsstudie zum Käferproblem mit der prognostischen Validität von Intelligenz beim Problemlösen. Die leicht veränderte Replikation ihrer ersten Studie, in der kein Zusammenhang zu beobachten war, konnte zeigen, dass je transparenter und dem analytischen Problemlösen ähnlicher die Problemstellung ist, umso mehr kann von leichten Zusammenhängen zwischen Problemlösekompetenz und Testintelligenz ausgegangen werden. Die Ergebnisse lassen vermuten, dass mit zunehmender Komplexität eines Problems weitere Faktoren als ausschließlich *reasoning*-Aspekte relevant werden. Diese Schlüsse stützen auch das Konzept von Problemlösen als operativer Intelligenz (Funke, 2003; Dörner, 1986), bzw. Problemlösen als intelligentem praktischen Handeln (Sternberg, 2003). Die Ergebnisse der ALL- und PISA-Studien zeigen ebenfalls, dass Problemlösen neben Intelligenz als eigenständiges Konstrukt zu identifizieren ist (Reeff, Zabal & Blech, 2006). Je mehr Anforderungen wie etwa planendes Handeln relevant werden, umso geringer ist die prognostische Validität der Testintelligenz. Wittmann und Süß (1999; auch Süß, Wittmann & Oberauer, 1991) operationalisieren Testintelligenz im Sinne Jägers Intelligenzstrukturmodell differenzierter und konnten vor allem die Faktoren Verarbeitungskapazität und Arbeitsgedächtniskapazität als Prädiktoren für die Problemlöseleistung feststellen. Der Faktor Bearbeitungsgeschwindigkeit erwies sich als eher unbedeutend (auch Sternberg 2005, 1981): Zwar bearbeiteten Personen mit niedrigerem Score Problemstellungen schneller als Personen mit höherem Testintelligenztestscore. Testintelligenter wandten die längere Zeit jedoch zum Enkodieren der Problemelemente auf. In der Problemlösung waren Testintelligenter letztlich schneller. Sternberg schloss daraus, dass Testintelligenter zwar mehr Zeit für globalere Planung (Analyse, Strategiewahl) aufwendeten, allerdings weniger Zeit für die lokale Planung benötigen (Ausführung der Problemlösestrategie, Evaluation, Sternberg, 2005).

Neben diesen Invarianten beim Problemlösen konnten Wittmann und Süß (1999; Süß, Wittmann & Oberauer, 1991) auch den Faktor Wissen als einflussreiche Größe identifizieren. Der Zusammenhang von Intelligenz und komplexem Problemlösen über den Faktor Wissen wird von Leutner (2002) als Elshout-Raaheim-Hypothese beschrieben (auch Veenman, Prins & Elshout, 2002): Intelligenz und Wissen hängen umgekehrt U-förmig zusammen, dies bedeutet, dass bei mittlerem Vorwissen ein maximaler Zusammenhang zwischen Intelligenz und Problemlöseleistung feststellbar sein sollte. Bei niedrigem oder hohem Wissen über die Problemsituation sollte die Korrelation sinken.

Insgesamt lässt sich damit feststellen, dass sich zwar bei testintelligenteren Personen schnellere Verarbeitungsgeschwindigkeit und längere Gedächtnisspannen, damit schnellere Enkodierung und *recalls* sowie differenziertere Strategieanwendung nachweisen ließen (auch Chi, Glaser, Rees, 1982). Darüber hinaus erweist sich jedoch der differentielle Wert der Testintelligenz für die Vorhersage problemlösenden Denkens als abhängig von der formalen Struktur des Problems und vorhandenem Wissen, Verarbeitungs- und Arbeitsgedächtniskapazität.

Die Rolle von Vorwissen und Expertise

Der Einbezug von Wissen und Expertise ermöglicht es, differentialpsychologisch Unterschiede zwischen Novizen und Experten, bzw. naiven und Master-Problemlösern zu beschreiben (Chi, 2006). Grundsätzlich unbestritten ist, dass als wohl beste Methode, domänenspezifische Problemlösekompetenzen zu steigern, diejenige gilt, entsprechende *skills* zu vermitteln. Expertise wird dementsprechend in erster Linie mit domänenspezifischem Wissen gleichgesetzt (Chi, Glaser & Farr, 1988; auch Ericsson & Smith, 1991; Mayer, 1992, 2003; Sternberg & Frensch, 1992). Die *knowledge-is-power*-Hypothese fasst diesen Zusammenhang von Problemlösen und Wissen zusammen. Zahlreiche Befunde verweisen darauf, dass domänenspezifisches Wissen fast 54% der Varianz von Problemlösekompetenz klärt (Hambrick & Engel, 2002). Auch hier stellt sich allerdings die Frage nach Wirkungszusammenhängen (für einen Überblick Chi, 2006; auch André & Fernand, 2008; Novick & Bassok, 2005; Ericsson, 2003b; Sternberg & Frensch, 1992; Chi, Glaser & Rees, 1982). Wie wird Wissen im Problemlöseprozess wirksam? Zahlreiche Studien legen dar, dass Experten nicht nur die besseren (de Groot, 1965; auch Segers, Dochy & de Corte, 1999), schnelleren und akkurateren (Klein, 1993, nach Chi, 2006) Lösungen als Novizen generieren. Vielmehr zeigt sich, dass Experten ProblemDetails, -muster und -merkmale sehr viel präziser erkennen als Novizen (Chi, Feltovich & Glaser, 1981). Chase und Simon (1973) stellen diesbezüglich fest, dass dies unter anderem mit einer elaborierteren Verknüpfung vorhandenen Wissens, das heißt, einer effektiveren Gruppierung von Einzelementen zu einer größeren sinnvolleren Einheit (*chunking*) bei Experten zu begründen ist. Mit Blick auf die Analyse von Problemsituationen sowie die anschließende Erstellung einer Problemrepräsentation konnte weiter gezeigt werden, dass Experten im Gegensatz zu Novizen Probleme eher qualitativ als quantitativ analysierten und dementsprechend genauere Problemrepräsentationen generieren konnten (Voss, Greene, Post, & Penner, 1983; Simon & Simon, 1978). Chi, Glaser und Rees (1982) berichten in Bezug auf die Phase der Problemanalyse davon, dass Experten im Gegensatz zu Novizen sowohl ihre eigenen Kenntnisse als auch mögliche Fehler und Schwierigkeiten im Problemlöseprozess besser einschätzen konnten. Mit Blick auf den weiteren Problemlöseprozess wissen zwar letztlich auch Experten nicht immer über die beste Strategie Bescheid (Lemaire & Siegler, 1995 nach Chi, 2006), sie verwenden jedoch sowohl die effektiveren als auch angemesseneren Strategien als Novizen, die überwiegend hypothesenprüfende Problemlösestrategien favorisieren (Lemaire & Siegler, 1995 nach Chi, 2006; Larkin, McDermott, Simon, & Simon, 1980). Catrambone (1996) berichtet davon, dass Experten im Gegensatz zu Novizen komplexe und nicht direkt lösbare Probleme eher in ihre unterschiedlichen Komponenten und Teilziele zerlegen, wobei sie während des Problemlöseprozesses eher strukturelevante Informationen und bedeutungstragende Verknüpfungen zwischen den Problemelementen fokussieren (Schiano, Cooper, Glaser & Zang, 1989; Chi, Glaser & Rees, 1982; Duncker, 1945). Einer effektiveren, schnelleren und präziseren Problemlöseleistung dienlich scheint Wissen dementsprechend auf zwei Arten zu sein:

- für Experten, indem sie eher auf Hintergrund- und Vorwissen zurückgreifen können,
- für Novizen, indem sie im Rahmen von Rückwärtsstrategien (bspw. Mittel-Ziel-Analysen) hypothesenprüfende Problemlösestrategien verfolgen können und über eine Verknüpfung von isolierten Einzelementen die Distanz zwischen Ausgangspunkt und Ziel zu reduzieren versuchen. Dabei fokussieren sie eher oberflächliche Probleminformationen, wobei sich mit zunehmender Expertise dieser Fokus verändert und die Aufmerksamkeit auf Tiefenaspekte und Strukturähnlichkeiten gelegt wird. Je mehr Expertise vorhanden ist, umso weniger spielen Oberflächenstrukturen eine Rolle (Novick & Bassok, 2005).

Letztlich gelingt Experten eine schnellere Problemlösung auch deshalb, da kognitive Prozesse automatisierter ablaufen (Alexander, 2003; Schneider, 1985, Chase & Simon, 1973; de Groot, 1965). Einfluss auf die effektivere und schnellere Problemlösung nimmt damit auch die ausgeprägtere Behaltensleistung sowie die elaboriertere Abspeicherung von sinnvollen Verknüpfungen (de Groot, 1965). Hinzu kommt, wie Chi (2006) berichtet, dass sich Experten im Gegensatz zu Novizen opportunistischer im Hinblick auf die Problemlösung zeigen und letztlich jedwede vorliegende und als sinnvoll erachtete Information zur Problemlösung nutzen. Funke (2003) berichtet auch davon, dass Experten beim komplexen Problemlösen manche Szenarien ausführlicher explorieren (Reither, 1981; *blindness of the specialists*, Funke, 2003).

Im Zusammenhang mit der Frage nach fächerübergreifenden Problemlösekompetenzen stellt sich letztlich die Frage, inwieweit Expertise das Lösen neuartiger Probleme unterstützt. Dahingehend kann Catrambone (1996) zeigen, dass vor allem diejenigen, die gelernt haben, ein Problem in Teilziele zu zerlegen, besser in der Lösung eines zwar ähnlich strukturierten, jedoch neuartigen Problems sind. Wie bereits Wertheimers Schulversuch zeigte, sind letztlich diejenigen Problemlöser zu besserem Transfer fähig, die ein gutes konzeptuelles Verständnis des Problems haben (auch Catrambone, 1996; Novick, 1988)

Allerdings hat Vorwissen nicht nur positive Effekte (Chi, 2006; Sternberg & Frensch, 1992; vgl. Gestalttheorie). Als problematisch erweist sich insbesondere ein eingeschränkter Fokus auf eine einzige Domäne. So überschätzen sich Experten gelegentlich in ihrer Kompetenz und übersehen mit ihrem Blick auf Details und Tiefenstrukturen Oberflächenmerkmale (Chi, 2006). Sie sind abhängig von Detailinformationen in ihren jeweiligen Domänen, das heißt, je mehr Kontextinformationen gegeben werden, umso besser lösen Experten zwar Probleme. Allerdings vertreten Experten oftmals die irriige Annahme, dass mehr Information zu akkurateren Lösungen führt. Zudem kommen Experten oftmals mit einer Veränderung der Situation, in der sie Expertise besitzen, weniger gut als Novizen zurecht (Sternberg & Frensch, 1992). Zudem konnte oftmals ein domänenspezifischer *bias* beobachtet werden. Das heißt, Hypothesen wurden ungeachtet ihrer Passung auf die Problemsituation eher dem Bereich entliehen, in dem Experten mehr Wissen haben (Hashem, Chi & Friedman, 2003).

Expertise ermöglicht dementsprechend Problemstrukturen präziser erkennen, Problemrepräsentationen adäquater entwickeln und Problemlösungen effizienter generieren zu können. Unterschiede zwischen Experten und Novizen zeigen sich dabei in drei Kategorien: Im Bereich der Aufgabenleistung, im Bereich der Strukturen der Wissensbasis sowie im Bereich der Wissensorganisation.

Problemlösen, affektive und metakognitive Faktoren

Emotionale Stimmungen können darauf Einfluss nehmen, ob ein Problem als solches angesehen und die Ausgangssituation als veränderbar, bzw. veränderungswürdig wahrgenommen wird, weiterhin können sie auch die Verarbeitungsmodi beeinflussen (Cohen, Henik & Mor, 2011; De Sousa, 2008; Schwarz & Skurnik, 2003; Isen, 1987). Der Einfluss affektiver Komponenten auf Problemlösen ist zwar zahlreich untersucht (bspw. Schwarz & Skurnik, 2003; Mayer, 2010, 1998), allerdings wird ihnen vor allem bei der Vermittlung von Problemlösekompetenzen immer noch wenig Beachtung geschenkt (Mayer, 1998). Daneben beeinflussen auch subjektive Überzeugungen als metakognitive Faktoren Lernen (Muis, Bendixen & Haerle, 2006; Bendixen & Rule, 2004). Auch ihnen wird bislang trotz ihrer handlungsleitenden Funktion zu wenig Beachtung geschenkt. Im Folgenden werden die Faktoren Interesse, Selbstwirksamkeit und epistemologische *beliefs* herausgegriffen, die, wie zahlreich belegt werden konnte, als handlungsregulierende Faktoren großen Einfluss auf Leistung nehmen (u.a. Pintrich & De Groot, 1990) und – wie im Fall der Selbstwirksamkeit, einen besseren Prädiktor für Leistung darstellen als intrinsische Motivation (Caprara, Vecchione, Alessandri, Gerbiono & Barbaranelli, 2011).

Mit Blick auf die Interaktion von Emotion und Problemlösen konnte gezeigt werden, dass ein emotional gefärbter Kontext Einfluss auf die Aufgabenbearbeitung nimmt (Blanchette & Leese, 2010, 2011). Weiterhin war festzustellen, dass die evozierte Betroffenheit nicht nur auf die Art der Bearbeitung des Szenarios Einfluss nimmt, das heißt, bspw. zu mehr Engagement bei solchen Problemen führt, die Betroffenheit wecken (Hesse, Spies & Lüer, 1983), auch die Art der Informationsverarbeitung wird beeinflusst. Es konnte gezeigt werden, dass bei positiven Emotionen eher einer intuitiv-heuristischen Vorgehensweise gefolgt wird, bei negativen Emotionen eher einer analytisch-sequenziellen, informationsgeleiteten (Schwarz & Skurnik, 2003). Positive Emotionen führen eher zu weniger Anstrengung bei der Detailerfassung, zu einer geringeren Detailorientierung und geringerer Effizienz bei der Problembearbeitung. Komplikationen in der Problembearbeitung führen jedoch auch bei positiven Emotionen zu einem Wechsel von einer *top-down* zu einer *bottom-up* Verarbeitung und zu einer Aufmerksamkeitsverengung auf Details (Wegner & Vallacher, 1986).

Während des Problemlöseprozesses stehen auch evaluative Aspekte im Zusammenhang mit Emotion: So konnte gezeigt werden, dass sich negative Stimmungen während des Problemlösens negativ auf die Einschätzung der Zielerreichung auswirken (Schwarz & Skurnik, 2003). Mit Blick auf die

Zielerreichung zeigte sich, dass positiv gestimmte Personen eher das erste zufriedenstellende Ergebnis annahmen, negative Stimmung hingegen führte dazu, weiter nach Optimierungsmöglichkeiten zu suchen und die erreichte Lösung eher als unbefriedigend anzusehen (Simon, 1967). Bereits im Zusammenhang mit der Instruktion wird Emotion geweckt: So konnte gezeigt werden, dass aufgabenbezogene Instruktionen aufgabenexplorativeres Verhalten nach sich ziehen – im Gegensatz zu selbstbezogenen Instruktionen (Funke, 2003).

Vor diesem Hintergrund wird der Einfluss von Emotion auch als *cognitive tuning* bezeichnet (Schwarz & Skurnik, 2003; Martin et. al., 1993). Insgesamt kann zusammengefasst werden, dass interessens- und motivationsbasierte Faktoren neben kognitiven die wichtigsten Prädiktoren darstellen, wobei insbesondere die Variablen Interesse und Selbstwirksamkeit bedeutsam sind (Schiefele, Krapp & Winteler, 1992; auch Mayer, 1998).

Interesse

Interesse als personenbezogenes Charakteristikum ist nicht nur als individuelles, sondern auch als situationales Interesse zu verstehen. Es bezieht sich auf die jeweilige Person-Situation Interaktion (Renninger, Hidi & Krapp, 1992). Schiefele, Krapp und Winteler (1992; Schiefele, Krapp & Schreyer, 1993) zeigen, dass Interesse mit einem mittlerem Zusammenhangsmaß von in etwa $r = .30$ Einfluss auf Leistung und Erfolg nimmt. Der Zusammenhang zwischen Leistung und Interesse ist als Schwellenwert zu interpretieren: Unterhalb dessen scheint Leistung eher von Fähigkeiten abzuhängen, oberhalb scheint der Einfluss motivationaler Faktoren zuzunehmen. Interessentheoretisch wird der Zusammenhang damit begründet, dass das am Fach gezeigte Interesse Verarbeitungstiefe und -breite beeinflusst (Schiefele, 1992), wobei ein hohes Interesse zu qualitativen, weniger zu quantitativen Wissensunterschieden führt (Schiefele, 1988). Relevant ist das Interessenskonstrukt auch mit Blick auf Transfer: So können Anand und Ross (1987; Ross et. al., 1985) zeigen, dass auf personalen Bezug ausgerichtete Aufgaben, das heißt, Aufgaben die mit Blick auf das Studienfach konzipiert waren, im Gegensatz zu studienferneren Problemstellungen mehr Transfer nach sich ziehen. Überflüssige, aber dennoch interessante Inhalte können allerdings auch den gegenteiligen Effekt nach sich ziehen: Harp und Mayer (1998) zeigen, dass interessante überflüssige Details vom Hauptgedanken der Problemstellung ablenken und Transfer verhindern.

Bislang ist allerdings noch ungeklärt, wie Interesse letztlich wirkt und in welche Komponenten das Konstrukt zerlegt werden kann (Mayer, 1989). Garner, Gillingham und White (1989, nach Mayer, 1989) zeigen, dass es für eine gute Leistung nicht ausreicht, interessante Einzelaspekte in Problemstellungen einfließen zu lassen. In Anlehnung an Kintsch (1980) unterscheiden Harp und Mayer (1997, 1998) zwischen emotionalem, effektorientiertem *enjoyment* und kognitivem, verständnisorientiertem Interesse. Sie zeigen, dass das Evozieren kognitiven Interesses Studierende in ihrem Lernen effizienter unterstützt als emotionales Interesse.

Subjektive Überzeugungen

Subjektive Überzeugung als metakognitive Faktoren beeinflussen Lernen (Muis, Bendixen & Haerle, 2006; Bendixen & Rule, 2004). Es wird angenommen, dass sie als selbstregulative Prozesse sowohl exekutive Routinen der Problemanalyse, -planung und -kontrolle beeinflussen und über kognitive Prozesse des Enkodierens, Vergleichens und Klassifizierens Einfluss auf Prozesse des Verstehens und des *conceptual change* nehmen. Darüber hinaus werden sie ebenso relevant im Hinblick auf handlungsleitende Prozesse, das heißt, die Wahl von Handlungsstrategien (für einen Überblick Hofer, 2008, 2001; auch Heigl & Thomas 2012; Bromme, Pieschl & Stahl, 2010; Buehl, 2008; Muis, Bendixen & Haerle, 2006; Schommer-Aikins, Duell & Hutter, 2005; Bendixen & Rule, 2004; Sandoval & Morisson, 2003; Baumert, Bos, Brockmann, Gruehn, Klieme, et. al., 2000; Schraw, Dunkle & Bendixen, 1995; Schwarzer, 1995; Sternberg, 1994). Überzeugungen bezüglich der Nützlichkeit dessen, was gelernt wird (Schoenfeld, 1983) oder bezüglich der Wirksamkeit des eigenen Handelns (Reid, 1987; Deci & Ryan, 1985) werden als Determinanten von *reasoning* gesehen. In zahlreichen Studien werden sie mit unterschiedlichen Aspekten des Lernens in Zusammenhang gebracht (Schommer-Aikins et. al., 2005; Bendixen & Rule, 2004; Schraw, Dunkle & Bendixen, 1995). Überzeugungen im Hinblick darauf, dass Intelligenz als unveränderbar oder Erfolg als zufällig angesehen wird, können Wissenserwerbsprozesse inhibieren. Überzeugungen, dass bspw. kognitive Fähigkeiten gesteigert werden können, können Wissenserwerb wiederum unterstützen (Sternberg, 1994). Sie beeinflussen Textverständnis, das Monitoring dieses Verständnisses oder auch den Gebrauch elaborierter Lernstrategien (Schommer-Aikins et. al., 2005). Das heißt, eine Bandbreite an subjektiven Überzeugungen steuert Verhalten und Leistung insofern, dass unsere Ideen, wie Situationen verändert werden können und bis zu welchem Grad wir selbst an diesen Veränderungsprozessen teilhaben können beeinflusst werden (Schwarzer, 1995).

Selbstwirksamkeit

Traue ich mir zu, eine bestimmte, gegebenenfalls auch neuartige Problemstellung zu bearbeiten? Selbstwirksamkeitserwartung als handlungsregulierender Faktor stellt einen besseren Prädiktor für Leistung als intrinsische Motivation dar (Pintrich & De Groot, 1990; für einen Überblick Caprara, Vecchione, Alessandri, Gerbiono & Barbaranelli, 2011). Selbstwirksamkeitserwartung als situative Erwartung bezeichnet die subjektive Überzeugung, schwierige Aufgaben aufgrund eigenen Handelns bewältigen zu können (Schwarzer & Jerusalem, 1999):

“Beliefs of personal efficacy constitute the key factor of human agency. If people believe they have no power to produce results, they will not attempt to make things happen. [...] Effective functioning requires both skills and the efficacy beliefs to use them. [Whereby] [p]erceived self-efficacy is concerned not with the number of skills you have, but with what you believe you can do with what you have under a variety of circumstances.“ (Bandura, 1997, S. 3, 36f.).

Zwei Aspekte werden dabei relevant: Die Erwartung, dass eine Handlung zu einem Ergebnis führt (*outcome expectancy*) und die Einschätzung des eigenen Handelns als erfolgreich im Hinblick auf das intendierte Ziel (*self-efficacy*) (Mayer, 1989): Wahrgenommene Selbstwirksamkeit wird dementsprechend definiert als:

"[...] Überzeugungen über diejenigen eigenen Fähigkeiten, die man benötigt, um eine bestimmte Handlung zu organisieren und auszuführen, um damit bestimmte Ziele zu erreichen." (Bandura, 1997, S. 3, Übers. nach Schwarzer & Jerusalem, 1999)

Der Grad der Überzeugung, dass die eigenen Fähigkeiten wirksam eingesetzt werden können, um eine bestimmte Aufgabe erfolgreich zu bearbeiten, beeinflusst Anstrengung und Ausdauer, die Kompetenzerwartung nimmt Einfluss auf die Situationsauswahl (Schwierigkeitsauswahl). Es liegen Evidenzen vor, dass Selbstwirksamkeit in Zusammenhang mit Strategiewahl (Pintrich & de Groot, 1990; Zimmerman & Martinez-Pons, 1990), Persistenz (Schunk, 1991) und Verarbeitungstiefe (Schunk & Hanson, 1985) steht.

Epistemologische Überzeugungen

Epistemologische Überzeugungen (für eine Begriffsdebatte Muis, 2007; Hofer, 2001) bezeichnen Annahmen über die Natur des Wissens und Lernens, bzw. den Wissenserwerb. Sie können als subjektive Theorien verstanden werden, die Menschen über Wissen und den Prozess des Wissenserwerbs im Allgemeinen sowie im Hinblick auf spezifische Domänen entwickelt haben (Hofer, 2008; Bendixen & Rule, 2004; Baumert, Bos, Brockmann, Gruehn, Klieme, et. al., 2000). Hofer definiert epistemologische Überzeugungen

"[...] as an identifiable set of dimensions of beliefs about knowledge and knowing, organized as theories, progressing in reasonable predictable directions, activated in context, operating both cognitively and metacognitively." (Hofer, 2005 nach Hofer 2008, S. 14)

Entsprechend dieser Definition will das Konstrukt die Multidimensionalität und Komplexität der subjektiven Vorstellungen und Annahmen über Wissen und Wissensentwicklung abbilden, das heißt Aspekte der Wissensgenese, -ontologie, -bedeutung, -rechtfertigung und -gültigkeit mit berücksichtigen. Epistemologische Überzeugungen werden als interkorreliert und domänenspezifisch verstanden (Bendixen & Rule, 2004), wobei die Person und ihre Auseinandersetzung mit Wissen und Wissenschaften im Zentrum steht (Hofer, 2008, 2004, 2001; Priemer, 2006; Moschner, Gruber & Studienstiftungsgruppe EPI, 2005; Baumert et. al., 2000).

Im Hinblick auf eine theoretische Umgrenzung des Konstrukts liegen zwar zahlreiche Untersuchungs- und Systematisierungsansätze vor (für einen Überblick u.a. Hofer, 2008; Bendixen & Rule, 2004). Allerdings konnte bis heute kein empirisch haltbares abschließendes Konstrukt vorgelegt werden (Bendixen & Rule, 2004). Im Allgemeinen wird in Weiterentwicklung von Modellen von Perry (1970), Belenky, Clinchy, Goldberger und Tarule (1986), Baxter Magolda (1992), Kuhn (1993) sowie King und Kitchener (1994) und Schommer (1990) von folgenden fünf Dimensionen ausgegangen:

- [1] Schnelles, bzw. schrittweises Lernen,
- [2] Lernen als angeboren oder veränderbar,
- [3] Wissen als isolierte Fakten oder komplexes System,
- [4] Sicherheit des Wissens,
- [5] Quelle des Wissens.

Obwohl die Anzahl der extrahierten Faktoren je nach zugrundeliegender Methode, diagnostischem Instrument und kulturellem Kontext variiert, verweisen Metaanalysen auf eine relativ stabile zweifaktorielle Struktur mit den Faktoren *Struktur des Wissens* (Dimensionen: Sicherheit des Wissens, Komplexität des Wissens) und *Struktur der Wissensentwicklung* (Dimensionen: Quelle des Wissens, Rechtfertigung des Wissens) (Hofer, 2008; für einen kritischen Überblick Buehl, 2008). Die Struktur des Wissens bezieht sich auf die Überzeugung bezüglich der Natur des Wissens, wohingegen die Struktur der Wissensentwicklung darauf abzielt, subjektive Überzeugungen hinsichtlich der Frage, wie Wissen gelernt wird, abzufragen. Diagnostische Instrumentarien basieren auf der Zweifaktorialität (bspw. EBI, Schraw et. al., 2002).

In Bezug auf Problemlösen zeigen zahlreiche Studien die prädikative Validität epistemologischer Überzeugungen (Schommer-Aikins, Duell & Hutter, 2005; Bendixen & Rule, 2004; Schraw, Dunkle & Bendixen, 1995). Ihre Bedeutsamkeit im Zusammenhang mit Problemlösen erhält die individuelle Epistemologie darüber, dass sie als selbstregulative Prozesse sowohl exekutive Routinen der Problemerkennung und -lösung als auch handlungsleitende Prozesse wie die Wahl von Lösungsstrategien beeinflussen (Hofer, 2008, 2001). Epistemologische Überzeugungen übernehmen instrumentelle Funktion, beeinflussen die Initiierung und Aufrechterhaltung von Lernprozessen und nehmen durch eine Vorstrukturierung der Inhalte Einfluss auf Lernergebnisse (Urhahne & Hopf, 2004; auch Cano & Cardell-Elawar, 2008; Hofer, 2008, 2000; Bendixen & Rule, 2004). Das heißt, sie regulieren die Annahmen des Lerners über Lerninhalte (Bromme, Pieschl & Stahl, 2010), Strategiegebrauch (Schommer, 1990; auch Urhahne & Hopf, 2004), kognitive Verarbeitung, konzeptuelles Lernen und das Verständnis des Erlernten (Hofer, 2008; Schraw et. al., 1995; Schommer, 1992). Insbesondere zeigte sich, dass naive epistemologische Überzeugungen wie die Überzeugung von der Unveränderlichkeit oder Einfachheit des Wissens mit weniger elaborierten Lernstrategien einhergehen (Buehl, 2008; Baumert et. al., 2000). Im Gegensatz dazu korrelieren reflektiertere epistemologische Überzeugungen mit höherem Interesse, einem ausgeprägteren Selbstkonzept und einer höheren Leistungsmotivation (Urhahne & Hopf, 2004).

Trotz zahlreicher Befunde zum Zusammenhang von *reasoning*, Problemlösen, Lernen und epistemologischen Überzeugungen sind die Korrelationen oftmals moderat. Einige Studien zeigen nur schwache Zusammenhänge (Hofer, 2008). Daneben können Sandoval und Morrisson (2003) zeigen, dass ein hoher Leistungszuwachs keineswegs immer auch mit einem Zuwachs an reflektierteren epistemologischen Überzeugungen einhergeht. Dementsprechend stellt sich auch die Frage, wovon die Entwicklung epistemologischer Überzeugungen abhängt und unter welchen Umständen Effekte epistemologischer Überzeugungen auf Problemlöseleistung zu beobachten sind (Muis, Bendixen & Haerle, 2006).

Die Herausforderung ist, dass der Einfluss epistemologischer Überzeugungen schwierig aufzudecken ist. Schommer-Aikins et. al. (2005) verweisen darauf, dass epistemologische Überzeugungen eine Art internalen Standard bilden. Als Konsequenz dieses Standards wird eine bestimmte Problemlösestrategie gewählt, die Anwendung dieser Strategie führt zu einer spezifischen Leistung. Das Konzept des internalen Standards ist dabei nicht auf kognitive Aspekte der Problemrepräsentation oder der Strategiewahl beschränkt. Vielmehr spielen auch individuelle Überzeugungen, wie bspw. Selbstwirksamkeit eine Rolle. Selbstwirksamkeit wird in diesem Zusammenhang als Faktor gesehen, der die Entwicklung elaborierter *beliefs* beeinflusst. Bendixen und Rule (2004) vermuten, dass "personal relevance", eine Kategorie in der sie vor allem auch das Konstrukt der Selbstwirksamkeit verorten, mit der Veränderung epistemologischer Überzeugungen verknüpft ist. Das heißt, Selbstwirksamkeit beeinflusst nicht nur den Gebrauch von Problemlösungsmethoden. Vielmehr beeinflusst Selbstwirksamkeit auch die Überzeugung, eigene Problemlösemethoden selbsttätig verändern zu können. Bendixen und Rule (2004) verweisen darauf, dass es bezüglich dieses Zusammenhangs von Selbstwirksamkeit, epistemologischen Überzeugungen und Leistung zwar zahlreiche Vermutungen gibt, allerdings bis heute wenige Studien.

Insgesamt ist festzustellen, dass im Hinblick auf die Wirkungsweise epistemologischer Überzeugungen und Zusammenhänge mit motivationalen und affektiven Variablen sowie Problemlösen zahlreiche Fragen offen sind (Buehl, 2008; Hofer, 2008, 2001; Kaartinen-Koutaniemi & Lindblom-Ylänne, 2008; Trautwein & Lüdtke, 2008; Schraw & Olofson, 2008). Dabei spielt auch die Frage der Domänenspezifität eine Rolle. Zwar werden epistemologische Überzeugungen als disziplinspezifisch definiert. Problematisch ist jedoch, dass disziplinäre Unterschiede bisher vernachlässigt wurden (Hofer, 2000). Es wird angenommen, dass Unterschiede zwischen Disziplinen integraler Bestandteil sind und mit zunehmender Expertise offensichtlicher werden (Kartinen-Koutaniemi & Lindblom-Ylänne, 2008). Begründet wird dies nicht nur mit Sozialisierungseffekten, sondern auch mit Effekten der Selbstselektion bei der Entscheidung für ein Studienfach (Trautwein & Lüdtke, 2008). Empirisch können diese Differenzen nachgewiesen werden (Kartinen-Koutaniemi & Lindblom-Ylänne, 2008; Trautwein & Lüdtke, 2008; Palmer & Marra, 2004). Unklar ist jedoch der Aspekt der Selbstselektion. So berichten unterschiedliche Studien verschiedene Evidenzen (Trautwein & Lüdtke, 2008; Lindblom-Ylänne, 1999; Ylijoki, 2000; Becher & Trowler, 2000), von einer

zunehmenden Differenzierung epistemologischer Überzeugungen im Laufe des Studiums berichten die Studien einheitlich.

Alters- und Geschlechtseffekte

Die Alterskorreliertheit kognitiver Kompetenzen wirft die Frage auf, inwieweit Zusammenhänge zwischen fächerübergreifendem Problemlösen und Alter zu erwarten sind. Daneben muss auch danach gefragt werden, inwieweit ein geschlechtsspezifischer *Bias* im Zusammenhang mit Problemlösen zu berücksichtigen ist.

In Bezug auf Einflüsse des Alters auf Problemlösen konnten McPherson (2002) sowie Hood und Deopere (2002) positive Effekte nachweisen. Diese beziehen sich auf die Entwicklung reflektierteren Denkens. Daneben konnten Alterseffekte auf *reasoning* nachgewiesen werden, die sich eher zuungunsten älterer Probanden auswirken (Salthouse, 2005).

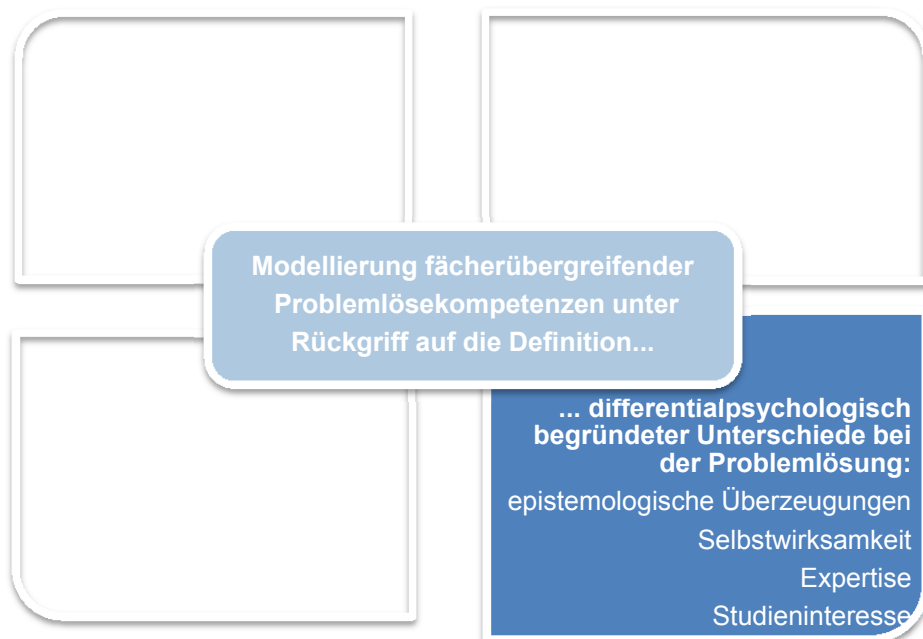
So zeigen sich leichte Alterseinflüsse auf Anweisungsverständnis (Salthouse, 1993), internale Verarbeitungsgeschwindigkeit (Salthouse, 1996) und -kapazität (Hambrick & Engle, 2002; Salthouse, 1991), Strategiegebrauch (Fristoe, Salthouse & Woodward, 1997) oder auch Arbeitsgedächtnis (für einen Überblick Salthouse, 2005; 1991). Weiterhin berichtet Salthouse vom Einfluss des Alters auf sowohl Testanweisung als auch internale mentale Operationen und entsprechende Verarbeitungsgeschwindigkeit, weist jedoch auch darauf hin, dass weit stärkere Alterseffekte vor allem auf Akkuratheit, Strategiegebrauch oder auch Verarbeitungskapazität zu berücksichtigen sind (Salthouse, 2005).

Als weitere Alterseffekte im Zusammenhang mit fächerübergreifendem Problemlösen konnten Einflüsse auf Informationsgenerierung, die Diskrimination von Beziehungen zwischen Variablen oder auf das Lösen von neuen und schwierigen Problemlöseaufgaben beobachtet werden. So konnte beispielsweise Hussy (1989) bei einem Vergleich von Problemlösekompetenzen bei acht bis sechzig Jährigen Operatoreffekte nachweisen. und berichtet davon, dass Problemlösestrategien bis zu einem Alter von 20 Jahren zunehmen, bei älteren Personen wiederum nimmt die Strategienutzung ab. Myers und Conner (1992) untersuchen, ob sich Alterseffekte beim Transfer zeigen und können signifikante Unterschiede zwischen jüngeren (19-20-Jährige) und älteren (30-59-Jährige) Probanden feststellen, zuungunsten der älteren im Hinblick auf deren Transferkompetenz (auch Cornelius, 1984; Denney, 1982, beide nach Salthouse, 1991). Bromley (1988 nach Salthouse, 2005) konnte zeigen, dass ältere Probanden bei der Problemlösung eher Oberflächenmerkmale fokussieren.

Die Frage nach geschlechtsspezifischen Unterschieden beim Problemlösen ist zwar vielfach untersucht, dennoch bis heute nur in Ansätzen geklärt: Die Befunde sind unterschiedlich, wobei sie nicht nur mit Blick auf die Effekte variieren, sondern vor allem auch mit Blick auf die den Studien zugrundeliegenden Problemstellungen (für einen Überblick Zu, 2007). Wird der Problemlöseprozess

betrachtet, so können Underwood, Underwood und Wood (2000) keine Geschlechtseffekte nachweisen, wobei der ihrer Studie zugrundeliegende Problemlösetest sprachliche Problemstellungen fokussiert. Auch McPherson (2002) kann in ihrer Studie zum Problemlösen bei Studierenden – die Grundlagen waren wissensarme *well-defined reasoning* Probleme – keine Geschlechtseffekte nachweisen. Die Ergebnisse von Underwood, Underwood und Wood (2000) zeigen weiter, dass mit Blick auf Problemlösen in Gruppenarbeit gleichgeschlechtliche Gruppen geschlechtsgemischten Gruppen überlegen waren. Harskamp, Ding und Suhre (2008) konnten dies replizieren. Geschlechtsunterschiede untersuchen auch Wittmann, Süß und Oberauer (1996). Ihre Studie bezieht sich auf komplexe Problemlöseszenarien. Sie finden in computerbasierten Problemlöseszenarien (Schneiderwerkstatt) Geschlechtsunterschiede zugunsten von Männern, diese Unterschiede können nur zum Teil durch Faktoren wie bspw. Computererfahrung erklärt werden. Dies steht im Einklang mit Befunden von Barbieri und Light (1992), die ebenfalls Unterschiede zuungunsten von Frauen finden. Auch ihrer Studie war eine computerbasierte Problemlöseaufgabe zugrundegelegt (nach Healy, Pozzi & Hoyles, 1995). Zahlreich untersucht sind Geschlechtseffekte vor allem in naturwissenschaftlichen Problemlösekontexten: Die häufig nachgewiesenen Effekte zugunsten von Männern (für einen Überblick Harskamp, Ding & Suhre, 2008; Zu, 2007) versucht Zu (2007) differenziert zu betrachten. So verweist er darauf, bei der Frage nach geschlechtsspezifischen Unterschieden Lernercharakteristika wie Strategiegebrauch, Einstellungseffekte wie Selbstwirksamkeit oder Selbstkonzept oder auch Sozialisationseffekte zu kontrollieren.

Abb. 4: Modell zur Beschreibung fächerübergreifender Problemlösekompetenzen



5.6 Die Frage der Messung fächerübergreifender Problemlösekompetenzen

Im Zentrum der vorangegangenen Kapitel stand, die Komplexität des Konstrukts fächerübergreifender Problemlösekompetenzen zu beschreiben. In Bezug auf das Modell fächerübergreifender Problemlösekompetenzen kann diesbezüglich festgehalten werden:

„[...] [T]o speak of a problem solving-ability tends to confuse the real issue.“ (Maier, 1939, S. 43)

Vor dem Hintergrund der beschriebenen Wirkungszusammenhänge und der Komplexität der Operationalisierung sind bestimmte Anforderungen an die Messung fächerübergreifender Problemlösekompetenzen zu stellen. Diese werden im Folgenden berichtet, bevor die für die vorliegende Studie leitenden Hypothesen vorgestellt werden.

Anforderungen an die Messung fächerübergreifender Problemlösekompetenzen

Unter Rückbezug auf die Anforderungen an die Kompetenzmessung sollte die Diagnostik fächerübergreifender Problemlösekompetenzen differenzierte Aussagen sowohl über das Kompetenzniveau als auch die Kompetenzstruktur erlauben und nicht darauf beschränkt bleiben, ausschließlich Lehreffekte im Sinne der Erreichung von Lehrzielen zu evaluieren. Im Sinne einer sowohl quantitativen als auch qualitativen Leistungsmessung gilt es, kognitive Entwicklungen nachzuzeichnen und differenzierte diagnostische Informationen und Indices zu generieren, um anhand einer Beschreibung möglicher Fehlerquellen Ableitungen zu deren künftiger Vermeidung zu treffen. Neben der empirischen Klärung des Konstrukts steht deshalb auch die Suche nach Fördermaßnahmen im Zentrum der Diagnostik fächerübergreifenden Problemlösens.

Mayer (1994) formuliert weitere Anforderungen an die Diagnostik des Problemlösens: Entsprechende testdiagnostische Verfahren sollten (1) den Bearbeiter mit neuartigen *non-routine* Problemen konfrontieren, die entsprechenden Testverfahren sollten (2) realistische Aufgabenstellungen beinhalten und (3) anspruchsvolle Denkoperationen (*higher-level problem solving*) implizieren. Grundsätzlich sollte das Kriterium der (4) Integration unterschiedlicher Anforderungen und Fähigkeiten erfüllt sein, wobei (5) kognitive Prozesse und Strategien fokussiert werden sollten und nicht nur Produkte des Problemlöseprozesses. Letztlich sollte das diagnostische Inventar (6) neben schriftlichen Aufgaben ohne Routinecharakter (7) idealiter auch Aufgaben miteinschließen, die sich in Abhängigkeit der Reaktionen des Bearbeiters verändern (*dynamic assessment*).

Baumert et. al. (o. J.) formulieren, dass ein Instrumentarium zur fächerübergreifenden Problemlösedagnostik Problemstellungen beinhalten sollte, die sich auf (8) inhalts- und situationsunabhängige Fähigkeiten beziehen, dementsprechend (9) in verschiedenen Domänen nachweisbar und (10) bei der Bewältigung komplexer Aufgaben von Bedeutung sind. Letztlich sollte das Instrumentarium Problemstellungen beinhalten, die Fähigkeiten erfordern die – wie Mayer (1994) ebenfalls festhält – (11) auf neuartige Problemstellungen übertragbar sind (Baumert et. al., o. J.).

Im Hinblick auf diese Anforderungen stellen Mayer (1994) und Reeffer (2006) fest (auch Klieme et. al., 2001a; Baumert et. al., o. J.), dass bislang noch keine Messinstrumente vorliegen, die diesen Kriterien entsprechen können.

Im Interesse der vorliegenden Studie liegen förder- und trainierbare kognitive Kompetenzen, die dem Lösen

- beruflich relevanter,
- fächerübergreifender,
- geschlossener Probleme

zugrundeliegen. Im Folgenden wird ein Überblick Instrumente vorstellen, die bislang zur Kompetenzmessung im Bereich des Hochschulassessments beziehungsweise des beruflichen Assessments verwendet werden und Ansatzpunkte für die Messung bei Studierenden liefern. Für einen detaillierten Überblick über die Bandbreite an diagnostischen Instrumenten zur Messung kognitiver *skills* und Problemlösekompetenzen im Allgemeinen wird auf Royer, Cisero und Carlo (1993), Pellegrino, Chudowsky und Glaser (2001) sowie Chi (2006) verwiesen, für einen kurzen Überblick über Verfahren zur Messung fächerübergreifender Problemlösekompetenzen auf Reeffer (o. J.), für einen einführenden Überblick über eignungsdiagnostische Verfahren auf Sarges und Scheffer (2008) oder auch Erpenbeck und Rosenstiel (2007b).

Überblick über diagnostische Instrumentarien zur Kompetenzmessung

Ein Blick auf die Vielzahl der Bereiche in deren Rahmen die Thematik des Problemlösens diskutiert wird, verdeutlicht, wie vielfältig die Zugangswege und diagnostischen Instrumentarien im Bereich des Problemlösens sind (für einen Überblick Chi, 2006; Royer, Cisero & Carlo, 1993; Dossey, Csapo, de Jong, Klieme & Vosniadou, o. J.). Diagnostische Zugänge zu Denk- und Problemlöseprozessen lassen sich grob klassifizieren in Selbstauskünfte, Verhaltensdaten und psychophysiologische Daten. Wird explizit die Diagnostik des Problemlösens als kognitive Kompetenz herausgegriffen, so zeigt sich, dass auch hier vielfältige Instrumentarien vorliegen: Es liegen Instrumente von der Analyse von Wissensstrukturen und der Wissensorganisation, über die Diagnostik von Problemlöserepräsentationen, metakognitiver Kompetenzen und Planungsprozessen beim Problemlösen bis hin zur quantitativen Überprüfung von domänenspezifischen Problemlöseleistungen. Grundlage für die Konstruktion des vorliegenden Messverfahrens sind vor allem diejenigen Verfahren, die im Bereich der hochschulischen und beruflichen Eignungs- und fächerübergreifenden Problemlösedagnostik vorliegen. Es soll ein empirisch haltbares paper-pencil-basiertes Instrumentarium entwickelt werden, das ökonomisch einsetzbar ist, weshalb im Folgenden ausschließlich auf empirisch bewährte Verfahren Bezug genommen wird. Auf die Vielzahl an informellen (bspw. Sullivan, 1994), technologiebasieren (bspw. Funke & Reuschenbach, 2011;

Bennett, Persky, Weiss & Jenkins, 2010; Vendlinsky, Baker & Niemi, 2008), domänenspezifischen (bspw. Jones & Davidson, 2007; Segers, Dochy & de Corte, 1999) oder auch dynamischen Verfahren (bspw. Dörner, 1989) zur Problemlösedagnostik wird an dieser Stelle deshalb nicht eingegangen.

Ein Blick auf die Assessmentverfahren im Hochschulsektor zeigt eine etablierte Verwendung spezifischer Diagnoseinstrumente zu Selektions- und Platzierungsverfahren. Eine Vielzahl formeller, informeller, domänenspezifischer und domänenunspezifischer diagnostischer Instrumentarien liegt insbesondere in den USA vor (bspw. *Scholastic Aptitude Test*; *American College Testing Assessment*; *Graduate Record Examination*; *Graduate Management Admission Test*; *California Critical Thinking Skills Test*; für einen Überblick bspw. University of Arizona, Online Ressource). Die Verfahren sind überwiegend auf Wissensdiagnostik ausgerichtet oder zielen darauf ab, kritisches Denken zu erfassen und zu trainieren (Vaughn, 2005). Die Intention ist, möglichst valide Prognosen über den Studienerfolg ableiten zu können, weshalb eine starke Orientierung an Intelligenztests vorhanden ist (bspw. SAT). Eine (prozessuale) Diagnostik von im Studium erworbenen Kompetenzen, die auf berufliche Anforderungen ausgerichtet ist, ist weniger etabliert.

In Deutschland werden eignungsdiagnostische Instrumente oder studienbegleitende Testverfahren zum kritischen Denken im Hochschulbereich bislang wenig verwendet. Eine Ausnahme bilden die Bereiche Medizin und Wirtschaftswissenschaften, teilweise basiert die Stipendiatenauswahl auf entsprechenden Verfahren. Das Testverfahren der Studienstiftung des deutschen Volkes umfasst beispielsweise unterschiedlich kontextualisierte Aufgaben zur Interpretation und Auswertung von Diagrammen und Grafiken, die Inbezugsetzung verschiedener Aussagen, das Erschließen von Analogien, Testaufgaben zum räumlichen Denken, Ableiten von Regeln und Gesetzmäßigkeiten sowie Textanalyseaufgaben. Der normierte Medizinertest (Deidesheimer Kreis, 1996; Trost, 1995) will Konzentration und Merkfähigkeit erfassen, umfasst Items zum analytischen Denken ebenso wie domänenspezifische Aufgaben zum naturwissenschaftlichen Denken. Insgesamt intendieren diese Verfahren, prognostisch valide Aussagen zum Studienerfolg abzuleiten, auch sie sehen sich der Tradition der Intelligenztests verpflichtet.

Weitere Ansatzpunkte bietet der Bereich der beruflichen Eignungsdiagnostik. Hier liegt eine große Bandbreite an Verfahren vor, die psychometrisch jedoch kritisch gesehen werden (Funke, 1993). Die Kritikpunkte fehlender Standardisierung, fehlender normativer Bezugssysteme und mangelnder Auswertungsobjektivität (Musch & Lieberei, 1997) wurden in den letzten Jahren zum Anlass genommen, zunehmend valide Instrumentarien zu entwickeln (bspw. Reuschenbach & Funke, 2000; Fennekels & D'Zouza, 1999). Verschiedene Typen von Fallstudien (*case-problem-methods*, *incident-methods*, *case-studies*) bieten Ansatzpunkte zur Konstruktion beruflich relevanter Problemsituationen (für einen Überblick Domsch, 1990). Meist müssen die Problemsituation sowie die Problemlösung definiert und entsprechende (Ziel)Konflikte erkannt werden. Mögliche Ursachen für problematische Ausgangslagen müssen identifiziert, Lösungsschritte und Alternativen definiert und im Hinblick auf

die Zielerreichung bewertet werden. Kritisch zu sehen sind jedoch die Einhaltung der Testgütekriterien sowie die testtheoretische Zuordnung der Items (Reuschenbach & Funke, 2000). Berufseignungsdiagnostik ist insgesamt eher ausgerichtet an einer eigenschaftstheoretischen Diagnostik beruflicher Kompetenzen und entspricht weniger einer Erfassung der im Rahmen des Problemlösens relevanten kognitiven Kompetenzen.

Demgegenüber können die im Rahmen der PISA-Studien konzipierten Testverfahren zum fächerübergreifenden Problemlösen überprüfte testtheoretische Ansatzpunkte liefern (für einen Überblick Klieme, Leutner & Wirth, 2005; OECD Testaufgaben PISA, Online Ressource). Fächerübergreifendes Problemlösen wurde hier sowohl als analytische, als auch als dynamische Problemlösekompetenz operationalisiert und im Rahmen der Aufgabentypen „Projektaufgaben“, „Raumfahrtspiel“, „Fehlersuche“ und „Analoges Problemlösen“ erfasst. Im Zentrum stand, Problemlösen in seiner Komplexität zu erfassen und kognitive Teilkompetenzen abzubilden. Dabei variierten die Aufgaben in den Kriterien Problemtypus, Komplexität und Problemkontext (für einen Überblick Leutner, Wirth, Klieme & Funke, 2005; Leutner, Klieme, Meyer & Wirth, 2004). Der Typus eines fächerübergreifenden Problems wurde darüber definiert, dass alle Problemstellungen zwar kontextualisiert, jedoch ohne Anwendung fachlichen Wissens zu lösen sind (Leutner, Wirth, Klieme & Funke, 2005). Im Folgenden werden die Aufgaben und Items kurz skizziert (für einen Überblick Baumert et. al., n.d.):

Die *Projektaufgaben* zum analytischen Problemlösen erfordern die Fähigkeiten, Informationen zu erfassen, diese zu bewerten, Entscheidungen abzuleiten, sowie planend und kombinatorisch zu denken. Die Probleme sind wohldefiniert, das Ziel ist, ein Projekt, das in mehrere Arbeitsschritte aufgegliedert ist, zu bearbeiten. Insgesamt 5 Teilaufgaben wurden konzipiert: Analyse einer Ausgangssituation im Rahmen der Teilaufgabe „Informationen sammeln“, Lösungsvorschläge entwickeln und bewerten im Rahmen der Teilaufgabe „Vorschläge sammeln“, Handlungen planen im Rahmen der Teilaufgabe „Temperatur messen“ sowie Ergebnisse bewerten und evaluieren im Rahmen der Teilaufgabe „Temperaturen vergleichen und neu einstellen“. Die Testaufgaben sind als geschlossene Multiple-Choice Aufgaben konzipiert.

Die Testaufgaben zum *Analogen Problemlösen* sind konzipiert als Planungs- und Ordnungsgeschichte, wobei aus einer Referenzaufgabe Lösungsschritte für die zu lösende Aufgabe abgeleitet werden müssen.

Die *Fehlersuche* impliziert, dass über den Aufbau eines mentalen Modells zur Funktionsweise eines technischen Gerätes (z.B. Luftpumpe) Ableitungen über vorhandene Fehler generiert werden müssen. Neben diesen geschlossenen Problemstellungen im paper-pencil-Format wurden auch computerbasierte Aufgaben zum *dynamischen Problemlösen* konzipiert („Virtuelles Labor“, „Ökologieplanspiel“, „Raumfahrtspiel“). Diese Probleme erfordern die Fähigkeiten, ein vorgegebenes System zu explorieren, Wissen über das System und seine Funktionsweise zu generieren und darauf

aufbauend Steuerungsaufgaben zu erfüllen. Der Problemraum wurde als endlicher Problemraum konzipiert. Relevante Informationen zur Funktionsweise müssen ausgewählt, Alternativen durchgespielt, Entscheidungen getroffen und korrigiert werden.

Alle konzipierten Aufgaben zum fächerübergreifenden Problemlösen erfordern die Fähigkeit, Zwischenschritte und Lösungsergebnisse zu bewerten und zu evaluieren (Leutner, Wirth, Klieme & Funke, 2005; Klieme, Hartig & Wirth, 2005).

Problemlösekompetenz im Rahmen der analytischen Problemstellungen wurde operationalisiert als Zahl der richtigen Antworten, im Rahmen der computerbasierten dynamischen Problemstellungen über die Effizienz der Exploration, das generierte Wissen sowie die anschließende Bewältigung der Steuerungsaufgaben (Leutner, Wirth, Klieme & Funke, 2005). Problemlösekompetenz wurde dementsprechend über die Qualität des Ergebnisses erfasst. Die Testergebnisse sollten die Teilkompetenzen messen, die zum Bewältigen des Problemlöseprozesses notwendig sind, das heißt die Fähigkeit, Informationen auszuwerten, zu verknüpfen und mit Blick auf bestimmte Ziele zu bündeln, um Handlungen zu planen, Entscheidungen zu fällen und zu evaluieren (Klieme, Hartig & Wirth, 2005). Die Schwierigkeit der Testaufgaben lag mit 0,68-0,33 im mittleren Bereich, die Reliabilitäten lagen mit Cronbachs $\alpha=.63-.79$ im akzeptablen Bereich. Das geschlossene Antwortformat erwies sich als angemessener als das offene (Klieme, Hartig & Wirth, 2005). Es zeigte sich, dass die schriftlichen Instrumente unproblematisch einsetzbar waren, die computerbasierten erheblich mehr Aufwand erforderten. Die Konfundierung mit Interesse und Geschlecht war insgesamt nicht signifikant, bzw. konnte im Subtest „Raumfahrtspiel“ über die Computererfahrung kontrolliert werden. Zusammenhänge mit Vorwissen wurden gefunden bei den „Projektaufgaben“ ($r=.19$), der „Fehlersuche“ ($r=.22$), dem „analogen Problemlösen“ ($r=.33$) und dem „Raumfahrtspiel“ ($r=.32-.35$). Die Adaption bisheriger Verfahren der Problemlösedagnostik auf den fächerübergreifenden Kontext scheint gelungen (Wirth, Leutner & Klieme, 2005). Die Schwierigkeit der Testitems war abhängig von der Anzahl der zu berücksichtigenden Variablen und der Zahl der Einzelentscheidungen, der Art ihrer Verknüpfung (kombinatorisch oder additiv) sowie der Anzahl der zur Informationsdarstellung verwendeten Formate (Grafiken, Texte, Zahlen) und der Offenheit des Systems.

6. Zusammenfassung

Im Folgenden wird die theoretische Verortung der Arbeit zusammengefasst.

Einleitend wurde dargestellt, dass das Konzept der fächerübergreifenden Kompetenzen ein zentrales Schlüsselement in der aktuellen Bildungsdebatte ist (bspw. Mayer, 2010; Rychen & Salganik, 2001). In diesem Zusammenhang stellt fächerübergreifendes Problemlösen eine wichtige Herausforderung für die Lehre an Hochschulen dar. Dabei wurde auch darauf verwiesen, dass bislang Kompetenzmodelle ebenso fehlen wie valide und reliable Messinstrumente. Zur theoretischen Verortung fächerübergreifender Problemlösekompetenzen wurde deshalb die Frage gestellt, wie kompetentes Verhalten definiert werden kann. Dabei wurde festgehalten, dass Kompetenz über die Bewältigung bestimmter Situationen beschrieben werden kann und erst aus der Realisation der Disposition erschließbar ist. Es wurden Leitlinien für den der Arbeit zugrundeliegenden funktionalen Kompetenzbegriff dargelegt, Kompetenzen werden demnach als erlernbare kontextspezifische Leistungsdispositionen verstanden, die sich auf Anforderungen in bestimmten Situationen beziehen, wobei die Breite der Domänen und Situationen variabel ist. Es wird einem rein kognitiven Kompetenzkonstrukt gefolgt. Kognitive Komponenten fächerübergreifender Problemlösekompetenz werden getrennt von motivationalen, volitionalen und sozialen Aspekten erfasst.

Der Messung von Kompetenzen sollte ein geeignetes psychometrisches Konzept zugrunde liegen. Kompetenzniveau- und Kompetenzstrukturmodelle bieten Ansätze für eine differenzierte Messung.

Die theoretische Erarbeitung kognitiver Kompetenzen fächerübergreifenden Problemlösens wurde ausgehend von den beiden konstituierenden Komponenten „fächerübergreifend“ und „Problemlösen“ geführt. Es wurde herausgearbeitet, dass im Zusammenhang mit der als Suchprozess zu beschreibenden Problemlösung im Rahmen von Analyse- und Syntheseprozessen vor allem kognitive Kompetenzen des logischen Schließens relevant werden. Ausgehend davon wurden kognitive Kompetenzen des logischen Schließens als Teilkompetenzen des Problemlösens dargestellt und die Teilfertigkeiten des deduktiven, induktiven sowie kausalen Schließens beschrieben.

Über die Frage der Disziplin(un)abhängigkeit wurde der Aspekt des *fächerübergreifenden* betrachtet und das Zusammenspiel zwischen disziplinbasiertem Kompetenzerwerb und kontextübergreifender Anwendung beleuchtet. In diesem Zusammenhang wurde die Frage nach Gelingensbedingungen des Transfers gestellt. Ausgehend von der Definition fächerübergreifender Problemlösekompetenz als Fähigkeit, neuartige Sachverhalte und Problemsituationen aufgrund des Erkennens bekannter Strukturen und Muster lösen zu können, wurde dargestellt, dass Effekte von Vorwissen auf neue Problemsituationen in Abhängigkeit bestimmter Schlüsselähnlichkeiten zwischen *target-* und Quellanalogie auftreten. Unähnlichkeiten erschweren selbst die Wirksamkeit von Hinweisen. Der Einfluss von Vorwissen zeigt sich dementsprechend insgesamt in Abhängigkeit der empfundenen form- und inhaltsbezogenen Nähe zwischen Ziel- und Basisanalogie sowie der gegebenen Hilfestellung.

Die Konzeption des Problems nimmt damit in entscheidendem Maße Einfluss auf seine Lösung, weshalb im Weiteren die Frage gestellt wurde, wie Probleme grundsätzlich unterschieden werden können. Die Schwierigkeit eines Problems ist letztlich die ausschlaggebende Determinante, um Problemlöseverhalten und -erfolg vorherzusagen zu können. Dabei sind die Kriterien Problemtypus, Repräsentationsform des Problems sowie individuelle Unterschiede in der Person des Problemlösers Anhaltspunkte zur Schwierigkeitssetzung. In besonderem Maße wird die Schwierigkeit eines Problems über seine Komplexität und die Art und Anzahl der Problemelemente variiert.

Um in der Person liegende Bedingungen des Problemlösens näher beschreiben zu können, wurden Wirkungszusammenhänge zwischen Problemlösen und Intelligenz, Expertise, affektiven Faktoren des Interesses und der Selbstwirksamkeit sowie selbstregulativen metakognitiven Komponenten wie epistemologischen Überzeugungen und letztlich Alter und Geschlecht skizziert.

Insgesamt wurde über diese theoretische Ausarbeitung der Anforderung Rechnung getragen, dass die Konzeptionalisierung fächerübergreifender Problemlösekompetenzen Aussagen über sowohl kompetenzstrukturbezogene Binnendifferenzierungen als auch Annahmen im Hinblick auf Kompetenzniveaus erlauben soll.

Abschließend wurde dargestellt, welche Anforderungen an die Diagnostik fächerübergreifender Problemlösekompetenzen zu stellen sind und welche diagnostischen Instrumentarien bereits vorliegen. Es zeigte sich, dass im Hochschulbereich in erster Linie die Frage der Studieneingangs-, weniger der Verlaufsdagnostik dominiert und Fragen der Kompetenzdiagnostik zur beruflichen Qualifikation weniger im Vordergrund stehen. Instrumente der beruflichen Eignungsdiagnostik, die außerhalb des Hochschulbereichs eingesetzt werden, sind im Hinblick auf testdiagnostische Gütekriterien kritisch zu sehen. Ein Instrumentarium, das den Anforderungen fächerübergreifender Problemlösekompetenzmessung entspricht, liegt mit den Testaufgaben im Rahmen der PISA-Studien (2003) vor. Problematisch an diesen ist, dass sie sich auf den schulischen Bereich beziehen, relevante Teilkomponenten des logischen Schließens im Rahmen des Problemlöseprozesses nicht berücksichtigen und insgesamt bedeutsame personbezogene Variablen, die Einfluss auf Problemlösen nehmen, nicht berücksichtigen. Dennoch dienen die Testaufgaben aufgrund ihrer Güte und ökonomischen Einsetzbarkeit als Ansatzpunkte zur Konstruktion des im Rahmen dieser Arbeit entwickelten Testverfahrens.

7. Forschungsfragen und Hypothesen

Die Arbeit fragt danach, wie sich fächerübergreifende Problemlösekompetenzen bei Studierenden empirisch abbilden und beschreiben lassen und inwieweit sich Wirkungszusammenhänge zwischen Aspekten des Problems und Faktoren, die in der Person des Problemlösers liegen, nachzeichnen lassen. Die Befunde sollen Anhaltspunkte für eine Förderung fächerübergreifender Problemlösekompetenzen im Hochschulbereich bieten.

Im Folgenden werden auf Basis der theoretischen Fundierung der Arbeit sowohl gerichtete als auch, bei ungenauer Evidenz, ungerichtete Hypothesen formuliert, um die der Arbeit zugrundeliegenden Forschungsfragen zu untersuchen. Die Hypothesen sind nach den für die vorliegende Arbeit leitenden Fragenkomplexen gegliedert:

Fragenkomplex I: Konstruktklärung dient der Überprüfung der empirischen Abbildung des theoretischen Modells fächerübergreifender Problemlösekompetenz. Es wird das der Studie zugrundeliegende theoretische Modell empirisch im Hinblick auf seine (Binnen-)Struktur geprüft.

Hypothesen

Der Prozess des Problemlösens kann über heuristische Modelle beschrieben werden (Kap. 5.2). Diese deskriptiven Modelle können über kognitive Teilkompetenzen messbar gemacht werden.

Daraus wird abgeleitet, dass ...

- **... sich über die Subtests kognitive Teilkompetenzen differenzieren lassen, die der Heuristik des Problemlöseprozesses zugeordnet werden können. (Hypothese 1).**

Fächerübergreifendes Problemlösen wird als mehrdimensionales kognitives Konstrukt definiert. Dabei werden unterschiedliche kognitive Teilkomponenten des logischen Schließens relevant (Kap. 5.2).

Daraus wird abgeleitet, dass ...

- **... sich über die Itemsets voneinander unterscheidbare Teilkompetenzen des deduktiven, induktiven und korrelativen Schließens sowie der Unterscheidung von Fakten und Meinung differenzieren lassen (Hypothese 2).**

Problemlösen und Intelligenz teilen umso mehr gemeinsame Varianz, je weniger komplex Probleme konstruiert sind und je weniger Planungs- und Steuerungsprozesse relevant werden (Kap. 5.5). Problemlösen und Testintelligenz teilen umso eher gemeinsame Varianz, je mehr die Testinstrumente Aufgaben des logischen Schlussfolgerns oder *reasoning*-Aufgaben umfassen, wie das Erkennen von Gesetzmäßigkeiten in numerischen oder symbolischen Darstellungen oder das Finden von Analogien.

Daraus wird abgeleitet, dass ...

- **... fächerübergreifende Problemlösekompetenzen und Testintelligenz in Abhängigkeit der zugrundeliegenden Anforderungsstruktur des Problems gemeinsame Varianz teilen (Hypothese 3).**

Abschließend soll untersucht werden, inwieweit sich Unterschiede zwischen Disziplinen beim fächerübergreifenden Problemlösen nachzeichnen lassen und inwieweit von einer Disziplingebundenheit fächerübergreifender Problemlösekompetenzen gesprochen werden kann. Auf Basis der berichteten Befunde (Kap. 5.3, 5.4) ist anzunehmen, dass Studierende, die bereits im Rahmen ihres Studiums mit ähnlichen Problemstellungen konfrontiert wurden, bessere Problemlösekompetenzen zeigen.

Daraus wird abgeleitet, dass ...

- **... sich Studierende der Sozialwissenschaften und Studierende der Geisteswissenschaften im Hinblick auf ihre Problemlösekompetenz unterscheiden. Dabei wird angenommen, dass Studierende der Sozialwissenschaften bessere Leistungen bei der Lösung von Problemen zeigen, die auf induktivem und kausalem Schließen basieren (Hypothese 4).**

Fragenkomplex II: Wirkungszusammenhänge zwischen personbezogenen Variablen und fächerübergreifenden Problemlösekompetenzen dient der Überprüfung von Zusammenhängen zwischen Expertise, epistemologischen Überzeugungen, Selbstwirksamkeit, Interesse, Intelligenz und fächerübergreifender Problemlöseleistung.

Hypothesen

Es wurde dargestellt, inwieweit motivationale und affektive Faktoren ebenso wie Geschlecht und Alter in Zusammenhang mit Problemlösen stehen (Kap. 5.5). Dabei wurde festgehalten, dass Interesse und Selbstwirksamkeit ebenso wie epistemologische Überzeugungen als handlungsregulierende Faktoren Einfluss auf Leistung nehmen. Expertise differenziert zwischen Novizen und Experten. Im Hinblick auf die Testfairness sollte ein geschlechtsspezifischer Bias ausgeschlossen werden.

Aus den bisherigen theoretischen und empirischen Befunden wird abgeleitet, dass ...

- **... sich Studierende mit einer höheren Selbstwirksamkeit und Studierende mit einer geringeren Selbstwirksamkeit im Hinblick auf ihre Problemlösekompetenz**

unterscheiden. Es wird angenommen, dass Studierende mit einer höheren Selbstwirksamkeit fächerübergreifende Probleme besser lösen können als Studierende mit einer geringen Selbstwirksamkeit. (Hypothese 5).

- **... sich Studierende mit einem höheren Studieninteresse und Studierende mit einem geringeren Studieninteresse im Hinblick auf ihre Problemlösekompetenz unterscheiden. Es wird angenommen, dass Studierende mit einem höheren Studieninteresse fächerübergreifende Probleme besser lösen können als Studierende mit einem geringeren Studieninteresse. (Hypothese 6).**
- **... sich Studierende mit reflektierteren epistemologischen Überzeugungen und Studierende mit weniger reflektierten epistemologischen Überzeugungen im Hinblick auf ihre Problemlösekompetenz unterscheiden. Es wird angenommen, dass Studierende mit reflektierteren epistemologischen Überzeugungen eine höhere Problemlösekompetenz haben als Studierende mit weniger reflektierten epistemologischen Überzeugungen (Hypothese 7).**
- **... Wechselwirkungen zwischen Selbstwirksamkeit, epistemologischen Überzeugungen und fächerübergreifender Problemlösekompetenz vorliegen. Es wird exploriert, inwieweit eine höhere Selbstwirksamkeit mit reflektierteren epistemologischen Überzeugungen und einer höheren fächerübergreifenden Problemlösekompetenz einhergeht (Hypothese 8).**

Fragenkomplex III: Fächerübergreifende Problemlösekompetenzen bei Studierenden beschreibt die Ausprägung fächerübergreifender Problemlösekompetenzen bei Studierenden. Insbesondere wird die Frage der Expertise exploriert, das heißt der Einfluss von Studienerfahrung und Studiendisziplin.

Hypothesen

Es wurde dargestellt, dass Expertise zwischen Novizen und Experten differenziert (Kap. 5.5). Darüber hinaus wurde dargestellt, dass kognitive Teilfähigkeiten des logischen Schließens in Zusammenhang stehen, bzw. vergleichbare kognitive Anforderungen stellen (Kap. 5.2). Schließlich wurde auch dargestellt, dass die Fähigkeit der Problemanalyse konstitutiv für die Problemlösung ist (Kap. 5.4).

Aus diesen Befunden wird abgeleitet, dass ...

- **... sich Studierende eines höheren Semesters und Studierende eines niedrigeren Semesters im Hinblick auf ihre Problemlösekompetenz unterscheiden. Es wird angenommen, dass Studierende mit einer höheren Expertise, das heißt Studierende eines höheren Semesters, eine höhere Problemlösekompetenz haben als Studierende eines niedrigeren Semesters (Hypothese 9).**
- **... diese Unterschiede vor allem bei schwierigen Aufgaben deutlich werden (Hypothese 10).**
- **... Unterschiede zwischen den Disziplinen beobachtbar sind. Es wird angenommen, dass Sozialwissenschaftsstudierende, bei denen eher davon ausgegangen wird, dass sie im Rahmen ihres Studiums überwiegend mit Fragestellungen konfrontiert sind, im Rahmen derer illegitime Kausalitäten vermieden werden müssen, diejenigen Problemstellungen besser lösen können, die Kompetenzen des induktiven und korrelativen Schließens erfordern“ (Hypothese 11).**
- **... kognitive Teilkomponenten fächerübergreifenden Problemlösens, die über ähnliche kognitive Anforderungen beschrieben werden können, auch empirisch in Zusammenhang stehen (Hypothese 12).**
- **... kognitive Kompetenzen zur Analyse von Problemsituationen (Subtest Situationsanalyse) in Zusammenhang stehen mit Teilkompetenzen der Bewertung und Handlungsplanung (Subtests Zielevaluation und Handlungsplanung) (Hypothese 13).**
- **... Studierende, die die Testitems des Subtests Situationsanalyse besser lösen, auch die Testitems der Subtests Zielevaluation und Handlungsplanung besser lösen (Hypothese 14).**

Abschließend wird ein Gesamtmodell fächerübergreifender Problemlösekompetenzen bei Studierenden erarbeitet.

Teil III: Testinstrument und methodisches Design der Studie

Im Folgenden wird die der Studie zugrundeliegende Methodik erläutert und das entwickelte Testinstrument beschrieben.

Einleitend wird der Rahmen für die formale Testkonstruktion abgesteckt. Es werden die Anforderungen, denen das konstruierte Verfahren Rechnung tragen muss, vorgestellt. Daran anschließend wird das Testverfahren beschrieben. In diesem Rahmen wird die Operationalisierung fächerübergreifender Problemlösekompetenzen sowohl auf Einzelitem- als auch Gesamtestebene unter Rückgriff auf die theoretische Verortung in Teil I und II der Arbeit beschrieben. Die Darstellung der Pilotierungs- und Hauptstichprobe sowie des zeitlichen Ablaufs der Vorstudie leiten über auf die Überprüfung der testtheoretischen Güte des Instrumentes und die Darstellung der die Hauptstudie begleitenden Testinstrumentarien.

Das nachstehende Schaubild zeigt das der Studie zugrundeliegende Design.

Grafik 3: Forschungsdesign der Studie

Entwicklung und Erprobung eines Instruments zur Diagnostik fächerübergreifender Problemlösekompetenzen bei Studierenden .

Forschungsdesign: Querschnitt

Die Hauptstichprobe umfasste

- ... Studierende im Erststudium
- ... Studierende unterschiedlicher Disziplinen (Geistes- und Sozialwissenschaften)
- ... Studierende unterschiedlicher Studiensemester (1., 3. und 6. Semester)

Eingesetzt wurden folgende diagnostischen Instrumente:

Diagnostik fächerübergreifender Problemlösekompetenz

Analyse der kognitiven Kompetenz fächerübergreifende Probleme lösen zu können

Kontrolle der Intelligenz

APM Set I
(dt. Bullheller & Häcker, 1998)

Personbezogene Kovariaten

Fragebogen zum Studieninteresse
(Krapp, Schiefele, Wild & Winterler, 1993)

Skalen zur Erfassung von Lehrer- und Schülermerkmalen , Selbstwirksamkeit
(Schwarzer & Jerusalem, 1999)

epistemologische Überzeugungen
(Urhahne und Hopf, 2004; eigenes Design)

8. Anforderungen an das Testverfahren

Im Folgenden werden Anforderungen an die formale Konstruktion des Testinstrumentes beschrieben. Für einen umfassenderen Einblick in Test- und Itemkonstruktion wird auf Haladyna und Rodriguez (2013), Bühner (2011), Moosbrugger und Kelava (2008), Downing und Haladyna (2006), Wilson (2005), Haladyna, Downing und Rodriguez (2002) oder auch Lienert und Raatz (1998) verwiesen.

8.1 Ziele des Testinstrumentes

Das Inventar soll zur Diagnostik kognitiver Kompetenzen eingesetzt werden, die im Rahmen des Problemlöseprozesses im fächerübergreifenden Kontext relevant werden. Das Instrument soll bei Studierenden eingesetzt werden. Ziel ist nicht nur die Erstellung und testtheoretische Überprüfung des Testverfahrens, sondern auch die Klärung des Konstrukts sowie die Analyse studentischer Kompetenzen. Dabei steht weniger eine ausführliche Individualdiagnostik im Zentrum, sondern vielmehr die Beschreibung der Kompetenzunterschiede. Die Analyse soll Aussagen erlauben, inwieweit bestimmte Gruppen differenziert werden können, das heißt, inwieweit Unterschiede zu erkennen sind und wie diese Unterschiede beschrieben werden können.

8.2 Zielstichprobe des Tests

Die Zielstichprobe des Testverfahrens ergibt sich aus der Fragestellung der Studie: Ausgangspunkt ist, dass Hochschulen vor der Herausforderung stehen, Studierenden fächerübergreifende Problemlösekompetenzen zu vermitteln. Daraus ergibt sich, dass sich das Testverfahren zunächst an Studierende aller Hochschulformen richtet. Das Testverfahren sollte demzufolge Studierende aller Studiendisziplinen ansprechen. Die Merkmale und Anforderungen des Testverfahrens orientieren sich an den kognitiven Anforderungen, die an Studierende vor Eintritt in das Berufsleben gestellt werden können. Eine testtheoretische Überprüfung der Schwierigkeit erfolgt im Rahmen der Pilotierung des Testverfahrens. Kognitive Anforderungen im Zusammenhang mit Problemlösen im Allgemeinen und fächerübergreifendem Problemlösen im Besonderen wurden in Teil I und II der Arbeit formuliert.

8.3 Formale Anforderungen an die Konstruktion des Verfahrens

Die Beschreibung der Ziele zeigt, dass das Testverfahren über verschiedene Studienphasen hinweg brauchbare Indices bezüglich Kompetenzniveau und Kompetenzstruktur generieren soll. Mit Blick auf die Frage nach Kompetenzstrukturen soll das Inventar das Konstrukt fächerübergreifenden Problemlösens über einen Globalindex erfassen, jedoch auch die Multidimensionalität des

Konstrukts über Einzelindices abbilden. Die Frage nach Kompetenzniveaus legt nahe, das Testverfahren als Powertest zu konzipieren. Die Differenzierung der Leistung erfolgt dabei über das Schwierigkeitsniveau der Items. Bei Testverfahren zur Erfassung des Kompetenzpotenzials sollte deshalb die Zeit zur Bewältigung der Testaufgaben bereit gestellt werden, die notwendig ist, um allen Testpersonen eine Bearbeitung bis zur letzten Aufgabe gewährleisten zu können. Nur so kann sichergestellt werden, dass eine unterschiedliche Anzahl nicht bearbeiteter Aufgaben nicht zu rechnerischen Problemen und Interpretationsschwierigkeiten führt (Rost, 2004). Der Einfluss der Verarbeitungsgeschwindigkeit soll gering gehalten werden (Bühner, 2011).

Die Zielstichprobe setzt voraus, dass das Testverfahren im Feldversuch einsetzbar ist: Testlänge, Testform und Testformat müssen ökonomisch und praktikabel sein. Das heißt, das Testverfahren sollte sowohl als Einzel- als auch als Gruppentest durchführbar sein und in paper-pencil Format vorliegen. Das Testverfahren sollte trotz der Komplexität des Konstrukts unmissverständlich bearbeitbar sein, eine für die Teilnehmenden annehmbare Bearbeitungszeit in Anspruch nehmen, geringen Materialverbrauch haben und objektiv und ökonomisch auswertbar und interpretierbar sein.

Möglichen Fehlerquellen in der Itembeantwortung sollte bereits über Testlayout und -instruktion entgegengewirkt werden (*Optimizing-Satisficing-Ansatz*, Krosnick, 1999): So kann die Motivation der Testbearbeitung bereits über ein ansprechendes Testformat sichergestellt und einer Beantwortung ohne große Anstrengung entgegengewirkt werden. Motivationsförderlich ist ebenso eine klare, präzise und unmissverständliche Testinstruktion. Eine entsprechende Instruktion stellt weiter sicher, dass der Untersuchungsgegenstand klar ist. Die Instruktion auf Subtestebene soll den Antwortmodus verständlich erläutern. Die Anonymität der Teilnehmer muss zugesichert und nur die nötigsten personalen Daten abgefragt werden.

8.4 Darstellung der begleitenden Information

Im Hinblick auf die formale Gestaltung des Testverfahrens muss die Frage der Darbietung der Informationen, die als Grundlage der Testaufgaben dient, geklärt werden. Gleiche Informationen werden abhängig von ihrer Darstellung unterschiedlich leicht verstanden. Eine Voraussetzung ist deshalb, dass die Informationen zur Bearbeitung der Testaufgaben einfach, kurz und klar verständlich dargestellt werden. Es soll ausgeschlossen werden, dass die Darstellung der begleitenden Information dann Einfluss auf das Testergebnis nimmt, wenn dies nicht intendiert ist. Im Folgenden werden die Konstruktionsprinzipien zur Darstellung der Informationen skizziert (Texte, Diagramme, Tabellen). Für einen umfassenderen Einblick wird auf Voss, Carstensen und Bos (2008), Artelt, Stanat, Schneider und Schiefele (2001), Kintsch (1998), Britton und Graesser (1996) sowie Schnotz und Bannert (1999) oder auch Wainer (1992) verwiesen. Für einen kurzen Überblick siehe auch Weidenmann (2006).

Das Lesen eines Textes zur Entnahme von Information ist keineswegs eine passive Rezeption der Informationen, vielmehr ist es eine aktive Rekonstruktion der Textbedeutung, wobei die Kompetenzen sowie das Vorwissen des Lesers und die objektive Textvorgabe interagieren (Artelt, Stanat, Schneider & Schiefele, 2001). Die Informationsaufnahme geschieht dementsprechend in Abhängigkeit des Verstehens des jeweiligen Textes, eine Konfundierung des Textverständnisses und der jeweiligen Lesekompetenz kann nie völlig ausgeschlossen werden.

Da die Testleistung nicht das Verständnis des Textes, sondern die Problemlösung und den aktiven Umgang mit sowie die logische Steuerung und Verknüpfung der gegebenen Einzelinformationen erfassen soll, sind bestimmte Konstruktionsprinzipien zu berücksichtigen, um den Einfluss der Lesekompetenz gering zu halten: Texte in Testaufgaben sollten kurz, einfach und unmissverständlich sein und wissensarm gestaltet werden. Das heißt, eine Informationsentnahme muss ohne Rückgriff auf Vorwissen erfolgen können. Die Bevorteilung bestimmter Studierendengruppen sollte ausgeschlossen sein, das heißt, das Aufgabenmaterial sollte thematisch entweder möglichst gleich nah oder gleich entfernt von allen Studierendengruppen sein.

Die Sätze sollten kurz und im Satzbau einfach sein, doppelte Verneinungen sollten vermieden werden, die Ausdrucksweise sollte der Zielgruppe bekannt und angemessen sein. Überflüssige Informationen konfundieren mit der Schwierigkeit der Aufgabe, der Text soll deshalb lediglich diejenigen Informationen vermitteln, die notwendig sind, es sei denn, die Schwierigkeit der Aufgabe soll variiert werden. Texte sollten gegliedert bzw. sequenziert sein, prägnant, anregend und kohärent, bzw. die Informationen nachvollziehbar logisch angeordnet (Weidenmann, 2006).

Diagramme, Grafiken oder Tabellen erfordern im Gegensatz zu realen Bildern beim Leser Übung und Erfahrung (Schnotz & Bannert, 1999): Die Effektivität der Analyse von Diagrammen hängt von der Interaktion der Darstellung und dem kognitiven System des Betrachters ab. Unterschieden wird zwischen subsemantischen (automatischer, vorwissensunabhängiger Prozess des Erkennens, der Diskrimination, Identifikation und Gruppierung der Komponenten) und semantischen Verarbeitungsprozessen. Ausschlaggebend für die Herstellung größerer Einheiten und den eigentlichen Verständnisprozess sind Galtsgesetze, über die bspw. ähnliche Aspekte zusammengefasst werden (z.B. Farben oder Formen). Das Verständnis eines Diagramms läuft letztlich im Rahmen dieser semantischen Verarbeitung ab und basiert auf der konzeptgeleiteten Analyse der wahrgenommenen Konfigurationen. Hier werden Vorwissen, Instruktionen (bspw. Angabe der Bearbeitungsreihenfolge etc.) oder auch kulturelle Einflüsse (bspw. Lesen von oben nach unten) relevant. Dabei hängt das Verständnis vor allem davon ab, inwieweit beim Betrachter bereits Schemata zur Analyse des Diagramms und zur Verknüpfung der gegebenen Informationen erlernt wurden, das heißt, inwieweit Vorwissen vorhanden ist. Wichtig ist deshalb, diese Wahrnehmungsprozesse konfliktfrei zu gestalten (deutlich unterscheidbare Farben, Formen, genügend große Schriften).

Gestaltprinzipien sind weiterhin Sparsamkeit, Passung von Grafik und Inhalt, Nähe von grafischer Darstellung und verbaler Information, ein logisch nachvollziehbarer Umgang mit Formen, Farben und Strukturen sowie klare Instruktionen und Verarbeitungshinweise. Auf irrelevante Effekte, die nicht der Information dienen, sollte verzichtet werden, die Komplexität des Diagramms sollte der Zielgruppe angepasst sein. Diagramme sollten insgesamt so gestaltet sein, dass ein möglichst einfaches, einheitliches, geschlossenes und symmetrisches Ganzes entsteht (für weitere Gestaltprinzipien auch Glug, 2009).

Wird die Verwendung von Tabellen betrachtet, so läuft die Verarbeitung entsprechend der Verarbeitung bei Grafiken ab (Wainer, 1992): Im Gegensatz zur Verwendung von Grafiken und Diagrammen kommt bei Tabellen jedoch hinzu, dass hier die Verwendung von Texten und Zahlen mathematische Grundfertigkeiten sowie grundlegende Lesefertigkeiten beim Betrachter voraussetzt. Insgesamt dienen Tabellen einer Zusammenfassung und Reduktion der Komplexität bestimmter Informationen. Dementsprechend steht bei der Erstellung im Vordergrund, ihren Komplexitätsgrad gering zu halten, das Herausarbeiten der relevanten Information einfach zu gestalten und die Tabellen sinnvoll und logisch aufzubauen. Im Vordergrund stehen, Tabellenstrukturen nach der Information zu sortieren, die für die Aufgabe relevant ist. Zahlen sollten gerundet dargestellt werden und auf die Darstellung überflüssiger Informationen sollte auch hier wieder verzichtet werden. Wichtige Informationen sollten hervorgehoben werden. Das Schwierigkeitslevel wird dabei weniger darüber variiert, dass mehr Informationen vom Betrachter extrahiert werden müssen, sondern eher darüber, dass verschiedene Arbeitsschritte anhand der Information bewältigt werden müssen.

8.5 Anforderungen an Testformat und Itemdesign

Kompetenzen lassen sich erst auf Basis einer Reihe von Einzelbeobachtungen bei unterschiedlichen Aufgaben, das heißt in variierenden Situationen abschätzen (Klieme & Hartig, 2007). Zur Diagnostik fächerübergreifender Problemlösekompetenzen bei Studierenden müssen dementsprechend eine Reihe verschiedener Situationen entwickelt werden, anhand derer von der Fertigkeit fächerübergreifende Probleme zu lösen auf die Fähigkeit zum fächerübergreifenden Problemlösen geschlossen werden kann. Relevante Prinzipien zur Konzeption von Multiple-Choice-Tests werden im Folgenden kurz dargestellt.

Leistungstests im Multiple-Choice Format sind zwar durch die Vorgabe von Antwortmöglichkeiten einschränkender als offenere Formate. Sie sind jedoch sowohl ökonomisch durchführbar als auch auswertbar und bieten die Möglichkeit einer standardisierten Durchführung und objektiven Auswertung. Die Reliabilität von Multiple-Choice Formaten ist dementsprechend höher als die offener Formate. Insgesamt, so stellen Duit, Häußler und Prenzel (2001) fest, werden Multiple-

Choice Formate in der Leistungsmessung bislang immer noch unterschätzt. Zwar ist die Ratewahrscheinlichkeit hoch, jedoch kann die Zuverlässigkeit des Formates über die Bereitstellung mehrerer Antwortmöglichkeiten oder auch die Hinzunahme mehrerer Aufgaben erhöht werden. Zur Konstruktion von Multiple-Choice Aufgaben gibt es unterschiedliche Subformate, wobei allen das gleiche Konstruktionsprinzip zugrunde liegt: Sie bestehen aus einem Aufgabenstamm, der die notwendigen Informationen enthält und die Aufgabenstellung beschreibt. Daraufhin werden verschiedene Antwortmöglichkeiten gegeben, die aus einer richtigen Antwort sowie proximalen und distalen Distraktoren bestehen. Aus den Vorgaben muss entweder die korrekte Antwort ausgewählt werden oder es wird die Richtigkeit des Aufgabenstammes abgefragt. Für die Analyse komplexer mentaler Prozesse wie logisches Schlussfolgern oder Problemlösen eignen sich Multiple-Choice-Formate sowohl im *constructed* als auch *selected response*-Format (Haladyna, 1994).

Neben der Frage des Formates sind weitere Konstruktionsprinzipien zu berücksichtigen:

Der Test muss gewährleisten, dass den Studierenden alle für die Bearbeitung notwendigen Informationen vorliegen. Die Beantwortung soll als ein Maß der fächerübergreifenden Problemlösekompetenz interpretiert werden können. Er soll nicht ihr Wissen zu einem bestimmten Themengebiet abprüfen.

Auf Einzelitemebene ist sicherzustellen, dass die Aufgaben unabhängig voneinander lösbar sind, das Testergebnis soll nicht aufgrund von Folgefehlern verfälscht werden. In Bezug auf die Itemschwierigkeit muss der Test an das Niveau angepasst sein, das von Studienabsolventen, bzw. Berufsanfängern erwartet werden kann. Allerdings muss das Testverfahren auch eine ausreichende Varianz ermöglichen und eine Differenzierung zwischen stärkeren und schwächeren Studierendengruppen erlauben. Die Itemformulierung sollte sich nicht auf die Schwierigkeit auswirken, der Schwierigkeitsgrad der Aufgabenbearbeitung sollte allein auf die Schwierigkeit der Aufgabe zurückzuführen sein. Fachsprachliche Formulierungen sollten deshalb vermieden werden, die Itemstämme sollten präzise und kurz formuliert sein, die Aufgabe sollte klar und unmissverständlich zu verstehen sein. Die einleitende Beschreibung und Aufforderung zur Aufgabenbearbeitung sollte einfach zu verstehen sein, eine zufällige Aufgabenlösung aufgrund des Nichtverständnisses der Aufgabe soll vermieden werden. Aufgrund der kohärenten Zielstichprobe (Abiturniveau, Alter zwischen 19 und 25 Jahren) müssen weitere formale Konstruktionsprinzipien (bspw. bestimmte Größe der Schrift, bestimmte Grafikgröße, nonverbales Instruktionsmaterial) nicht expliziter berücksichtigt werden. Eine gute Lesbarkeit sollte jedoch gewährleistet sein.

8.6 Reihenfolgeeffekte und Itemfolge

Das Testverfahrens muss logische Abhängigkeiten und den Einfluss von Reihenfolge- und Positionseffekten bei der Itemabfolge berücksichtigen (für einen Überblick Krosnick, 1999; auch

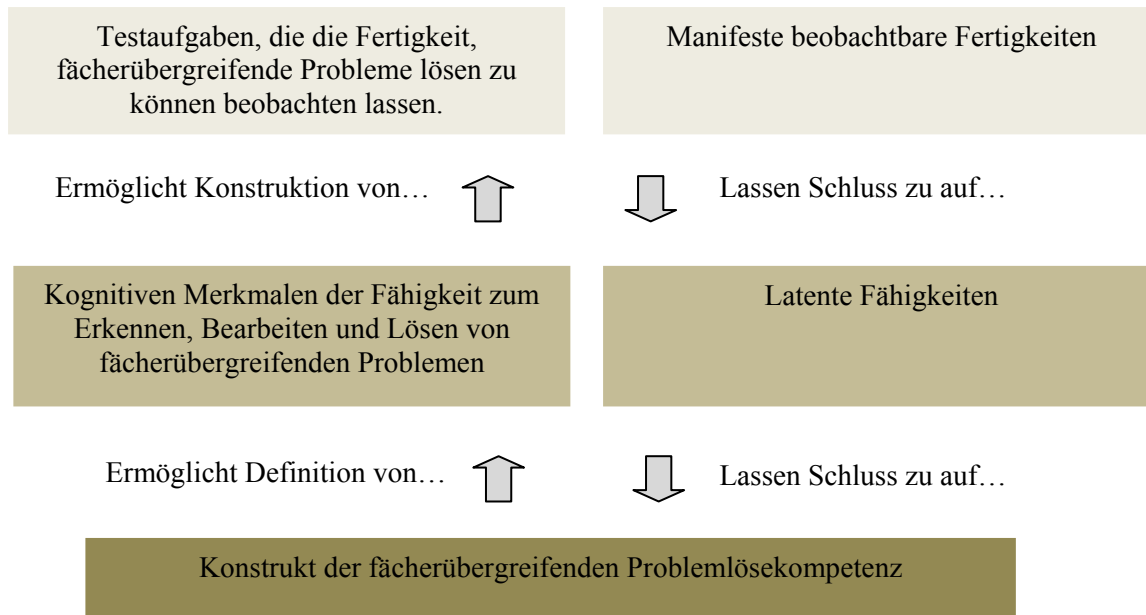
Bühner, 2011; Rost, 2004; Haladyna, Downing & Rodriguez, 2002): Das bedeutet, aufgrund der Reihung der Items sollte die Fähigkeit der Testperson weder über- noch unterschätzt werden, da bspw. zu schwierige Aufgaben am Anfang stehen und eine zu lange Bearbeitungszeit dieser Aufgaben die nachfolgende Bearbeitung leichterer Items aufgrund von Zeitmangel nicht mehr erlaubt. Andererseits sollte darauf geachtet werden, dass bei einer Darbietung der Items nach aufsteigender Schwierigkeit schwierigere Items aufgrund von Übungseinflüssen eventuell leichter zu lösen sind. Auch könnten bspw. thematisch weniger interessante Items zu einer Abnahme der Motivation führen. Zudem musste ausgeschlossen werden, dass die Beantwortung einzelner Items die Antwort auf ein folgendes Item erleichtert oder erschwert (Rost, 2004). Auch Positioneffekte, als Veränderungen der Schwierigkeit oder anderer Merkmale eines Items infolge seiner Platzierung wurden versucht zu vermeiden. Über eine ausführliche Instruktion sowie Beispielitems wurde Anfangsschwierigkeiten in der Aufgabenbearbeitung entgegenzuwirken versucht. Zudem wurde versucht, Effekten gegen Testende wie bspw. Ermüdung, Absinken der Testmotivation oder Zeitmangel durch die Testkonstruktion entgegenzuwirken. Berücksichtigung fand ebenso, dass Positioneffekte auch bei der Positionierung der richtigen und falschen Antworten innerhalb des Antwortformats einzelner Items auftreten können. Für die Testkonstruktion und Itemabfolge war deshalb insgesamt leitend, dass Items keine logische Abhängigkeit aufweisen dürfen, weshalb sowohl die genannten Konstruktionsprinzipien berücksichtigt wurden, als auch eine zufällige Itemfolge gewählt wurde.

8.7 Abgrenzung des vorliegenden Testverfahrens

Das Testverfahren setzt sich zum Ziel, das Konzept, das im Rahmen der PISA-Studien zum fächerübergreifenden Problemlösen für den Schulbereich entwickelt wurde aufzugreifen und weiterzuentwickeln. Das vorliegende Verfahren grenzt sich von auf Selektion und Platzierung ausgerichteten Wissens- und Leistungstests im Hochschulbereich sowie von Intelligenztests ab. Es soll ein Verfahren entwickelt werden, das auf den Berufseinstieg ausgerichtet ist. Allerdings ist es von im beruflichen Sektor vorhandenen weniger validen Verfahren abzugrenzen. Das Testverfahren sollte es ermöglichen, eine möglichst valide Möglichkeit zu bieten, Stärken und Schwächen im fächerübergreifenden Problemlösen bei Studierenden aufzudecken und Ableitungen für die Lehre zu ermöglichen.

9. Beschreibung des Testverfahrens: Das formale Testdesign

Die manifesten und beobachtbaren Fertigkeiten sollen dementsprechend Rückschlüsse auf zugrunde liegende latente Fähigkeiten, das heißt, auf die fächerübergreifende Problemlösekompetenz erlauben.

Grafik 4: Operationalisierung der Testaufgaben**9.1 Die formale Test- und Itemstruktur**

Das Testverfahren soll Kompetenzniveaus und Kompetenzstrukturen messen. Die Teststruktur ergibt sich aus diesem Messmodell, aus einer ausreichend großen Grundgesamtheit an Items sowie der Berücksichtigung der angemessenen zeitlichen Vorgabe. Jedes Item sollte eindeutig interpretierbar sein. Es muss mit einer 50% -Ausfallquote von Items gerechnet werden (Wilson, 2005), weshalb zunächst die doppelte Itemanzahl entwickelt wurde. Aufgrund verschiedener Aspekte (kognitive Anforderung der Items, schwierige und leichtere Items, Abbildung der Komplexität des Konstrukts über verschiedene Teilkompetenzen, Testökonomie) konnte beim Testverfahren, das dem Haupttest zugrunde liegt, die Anforderung von mindestens fünf Items pro zu messender Fertigkeit nicht eingehalten werden. Auf Ebene der Subskalen bilden nach der Pilotierung deshalb mindestens zwei Items ein Itemset, das heißt, eine kognitive Fertigkeit wird über zwei Items gemessen. Für die Pilotierung wurden mindestens vier Items pro zu messender Fertigkeit entwickelt. Insgesamt ergab sich mit den Subskalen des Testverfahrens ein 27 Items umfassendes erstes Testheft für die Pilotierungsphase (Subtest Situationsanalyse: 19 Items; Subtest Zielevaluation: 4 Items; Subtest Handlungsplanung: 4 Items). Dieser Umfang nahm eine mehr als 90 minütige Bearbeitungszeit in Anspruch, das Ziel der Pilotierung war deshalb nicht nur die Auswahl geeigneter Items, sondern auch die Reduktion auf wenige, möglichst aussagekräftige Items. Die Bearbeitungszeit für die Hauptstudie konnte auf 60 Minuten reduziert werden (13 Items: Subtest Situationsanalyse: 9 Items; Subtest Zielevaluation: 2 Items; Subtest Handlungsplanung: 2 Items).

Die Informationen, die zur Bearbeitung notwendig war, wurden auf Subtestebene und in den Itemstämmen einfach und logisch nachvollziehbar dargestellt. Das begleitende Material wurde

nach Testaufgaben gereiht dargeboten: Texte, Graphiken und Tabellen sind in Begleitheften dem Aufgabenmaterial beigelegt und nach Subtests und Itemfolgen geordnet. Insgesamt wurden zwei Hefte mit begleitenden Testmaterialien erstellt, für Subtest Handlungsplanung wurden die Materialien im Testheft dargeboten.

Für das Testverfahren wurde ein hypothetisches berufliches Setting entwickelt, in das sich die Studierenden hineinversetzen sollten. Über ein ansprechendes und reales Tätigkeitsszenario sollte gewährleistet werden, dass ein anregendes, im späteren Beruf mögliches Testszenario vorgegeben war und die Analyse fächerübergreifender Problemlösekompetenzen über die Bearbeitung ökonomisch valider Problemszenarien möglich sein würde. Es wurde versucht, ein berufliches Szenario zu entwickeln, das zwar von allen Studierendengruppen als spätere berufliche Tätigkeit vorstellbar, jedoch auch allen Studierenden gleichermaßen unbekannt sein sollte. Durch die Vorgabe eines übergeordneten Szenarios wurde vermieden, dass die Aufgabenstruktur durch unterschiedliche thematische Szenarien auf Subtest- oder Itemebene möglicherweise unübersichtlich wurde und unterschiedlich motivierend wirken könnte. Das Gesamtszenario wird im Rahmen der Testinstruktion vorgestellt, alle Testpersonen werden auf den gleichen Wissensstand gebracht, die Studierenden werden weiterhin über das Ziel der Studie und über die Freiwilligkeit der Teilnahme aufgeklärt. Auf Subtestebene wird in das spezifische Aufgabenszenario eingeführt. Damit ist gewährleistet, dass das übergeordnete Testszenario durch das Testverfahren führt und dem Test einen kohärenten Rahmen gibt, die einzelnen Aufgaben jedoch unabhängig voneinander konstruiert sind.

Den Items vorangestellt ist jeweils ein Itemstamm, der die Aufgabenstellung enthält.

Der Itemstamm führt in einen Themenbereich ein und legt die jeweils zu bearbeitende Aufgabe dar. Die Aufgaben folgen dabei folgenden Konstruktionsprinzipien: Die Problemstellungen sind einem der zu messenden kognitiven Fertigungsbereiche zuzuordnen. Zur Beantwortung der Items wurde ein gebundenes Format gewählt: Multiple-Choice für die Subtests 1 und 2 sowie das Umordnungsformat für Subtest 3. Die Subtests 1 und 2 bestehen aus einer Form von Richtig-Falsch Items. Sie können von den Testpersonen leicht verstanden und trotz der Komplexität der Aufgaben schnell bearbeitet werden. Um Zufallslösungen zu reduzieren, wurden insgesamt vier Antwortmöglichkeiten angeboten. Subtest 3 bietet für jedes der entwickelten Items eine Auswahl an fünf Handlungsschritten dar, die in die richtige Reihenfolge gebracht werden müssen. Das Antwortformat richtet sich nach der zu testenden Fähigkeit, das heißt, es misst die Anwendung der vorliegenden Informationen auf die vorliegende Problemstellung, wobei die Anwendung der Information über die Manipulation der gegebenen Informationen im Sinne der zu testenden kognitiven Operation erfolgt. Das Antwortformat besteht aus einer richtigen Antwort und verschiedenen Distraktoren. Distraktoren sollen die Testperson von der richtigen Lösung ablenken (Bühner, 2011), sie müssen deshalb sowohl plausibel erscheinen, dürfen jedoch weder zu schwer,

noch zu leicht sein. Ein guter Distraktor muss für Testpersonen mit geringer Kompetenz attraktiv und für Testpersonen mit hoher Kompetenz erkennbar sein (Lienert & Raatz, 1998).

9.2 Operationalisierung fächerübergreifender Problemlösekompetenzen: Test- und Itemdesign

Die Testaufgaben basieren auf der Definition von fächerübergreifender Problemlösekompetenz als kognitiver Kompetenz zur Analyse und Synthese gegebener Problemelemente im fächerübergreifenden Kontext. Orientierung zur Konzeption der Subtests gaben deskriptive Modelle des Problemlösens die die zentralen Problemlösephasen beschreiben. Das Setting des Testverfahrens wird über den Aspekt des fächerübergreifenden bestimmt: Dem Testverfahren liegt ein bedeutsamer beruflicher Alltagskontext zugrunde, wobei sich die fächerübergreifenden Problemstellungen jedoch dadurch auszeichnen, dass sie ohne weiteres Wissen gelöst werden können. Es wird davon ausgegangen, dass die kognitiven Kompetenzen, auf die bei der Aufgabenbearbeitung zurückgegriffen wird, insofern fachübergreifenden Charakter haben, dass sie situations- und inhaltsunabhängig relevant werden und bei der Lösung neuartiger, unbekannter Probleme sowie bei der Bearbeitung komplexer und ganzheitlicher Anforderungen erforderlich sind. Über die Komplexität der Items wurde die Schwierigkeit definiert. Die Komplexität wird manipuliert über

- die Art und Anzahl der kognitiven Operation,
- die Anzahl der zu manipulierenden Variablen und
- die Modalität der Repräsentation.

Die Testitems haben dementsprechend eine bestimmte **kognitive Anforderungsstruktur** (kognitiv höhere Anforderungen schlussfolgernden Denkens, Relevanz der Problemanalyse, Erkennen relevanter/einschränkender Bedingungen, planvolles Vorgehen) sowie eine bestimmte **konstruktionsbezogene Aufforderungsstruktur** (sprachlogische Komplexität, Realitätsbezug, nicht eindeutig einem Fachgebiet zuzuordnen, kein dezidiertes curriculares Wissen) und unterscheiden sich in ihrer **Schwierigkeit** (Manipulation des Problemtypus, Manipulation der Repräsentationsform).

Ausgehend von dieser funktionalen Definition wurden drei Subtests mit folgenden Testitems entwickelt:

- **Testaufgaben zur Messung kognitiver Kompetenzen, die bei der Analyse der Problemsituation bedeutsam werden (Subtest 1: Situationsanalyse):** Hier wurden Testaufgaben konstruiert, anhand derer auf die kognitive Fähigkeit zur Problemanalyse im fächerübergreifenden Kontext geschlossen werden kann. Beobachtet werden kognitive Fertigkeiten des induktiven, deduktiven und kausalen Schlussfolgerns. Darüber hinaus

wurden Aufgaben konstruiert, bei denen zwischen objektiven (Fakten) und subjektiven (Meinung) Informationen unterschieden werden muss.

Itemsets wurden gebildet zum

- induktiven Schlussfolgern (2 Items)
- deduktiven Schlussfolgern (3 Items)
- Unterscheiden von Korrelation und Kausalität (2 Items)
- Unterscheiden von objektiven und subjektiven Informationen (2 Items).

Die Schwierigkeit wurde definiert über die Art der kognitiven Operation, die Variation des Stimulusmaterials sowie beim Itemset zum deduktiven Schließen über die Anzahl der zu manipulierenden Variablen, weshalb das Itemset zum deduktiven Schlussfolgern aus drei Items besteht.

- **Testaufgaben zur Beobachtung kognitiver Kompetenzen, die bei der Handlungsplanung bedeutsam werden (Subtest 3: Handlungsplanung):** Im Rahmen dieses Subtests wurden Testaufgaben konstruiert, die kognitive Fähigkeiten beobachtbar machen sollen, die bei der Zerlegung eines komplexeren Problems in Handlungsschritte virulent werden und die auf die Lösungsplanung beim fächerübergreifenden Problemlösen schließen lassen.

Itemsets wurden gebildet zur Dekomposition und Planung, das heißt zur

- Zerlegung komplexer Situationen und
- Reihung von Handlungsschritten.

Die Schwierigkeit wurde definiert über die Art der kognitiven Operation und die Art der Distraktoren (proximal oder distal).

- **Testaufgaben zur Beobachtung kognitiver Kompetenzen, die bei der Evaluation von Zielen bedeutsam werden (Subtest 2: Zielevaluation):** Hier wurden Testaufgaben entwickelt, über die auf kognitive Fähigkeiten geschlossen werden kann, die während der Lösungsbewertung relevant werden. Über diese Testaufgaben soll auf die Kompetenz zur Einordnung von Zielen sowie das Erkennen von Zielkonflikten geschlossen werden.

Es wurde ein Itemset gebildet, um

- vor dem Hintergrund einer Problemsituation Problemlösungen bewerten und einordnen, Zielkonflikte erkennen und Lösungsmöglichkeiten einordnen zu können.

Die Schwierigkeit wurde variiert über die Art der kognitiven Operation und die Modalität des Stimulusmaterials (Statistik, textbasiert).

Die Subtests zielen darauf ab, zwischen denjenigen kognitiven Kompetenzen zu differenzieren, die während der Situationsanalyse, der Handlungsplanung sowie der Zielbewertung beim Problemlösen im fächerübergreifenden Kontext relevant werden. Dabei werden bei allen Aufgaben kognitive Operationen des Suchens, Erkennens, Verknüpfens und Evaluierens von Informationen relevant. Im Rahmen der Ergebnisdarstellung wird nochmals dezidiert auf die Konzeption eines jeden einzelnen Items eingegangen.

Testszenario und Itembeispiele

Im Folgenden wird das dem Haupttest zugrundeliegende Testszenario dargestellt (Testverfahren der Pilotierung und der Hauptstudie siehe Anhang). Die Aufgabenstämme der Subtests werden dargestellt, die Itemsets und einzelnen Aufgaben werden anhand konkreter Itembeispiele verdeutlicht. Im Rahmen der Darstellung der Ergebnisse wird bei der Beschreibung und beim Vergleich der Schwierigkeiten der Studierenden nochmals dezidiert auf die kognitive Anforderungsstruktur der einzelnen Items, das heißt die Konzeption der einzelnen Problemelemente, eingegangen. Aus Gründen der Kürze wurde an dieser Stelle darauf verzichtet.

Einleitend wird den Testpersonen die Intention der Studie dargelegt. In dieser Darstellung werden allgemeine Hinweis zur Testbearbeitung gegeben, die Testpersonen werden gebeten, nach einem vorgegebenen Schema ein Codewort zu vergeben, sodass die einzelnen Teile der Studie, die an zwei Testtagen durchgeführt wurde, ohne Anonymitätsverletzungen wieder zugeordnet werden können. Auf der danach folgenden Seite werden die Testpersonen gebeten, sich in das skizzierte Szenario zu versetzen:

Wir befinden uns im Jahr 2011. Seit kurzem sind Sie Referent/Referentin des amtierenden Bürgermeisters der Stadt Beispielingen, Hermann Eckerle (erste Legislaturperiode: 2006-2011, Mitglied der Partei GUP, Gemeinschaftliche Unabhängige Partei). In wenigen Monaten stehen Neuwahlen an und natürlich würde Hermann Eckerle gerne ein weiteres Mal im Amt bestätigt werden. Dies könnte schwierig werden, denn mit Dr. Johann Igerle (Mitglied der SFP, Soziale Freie Partei) hat die Opposition einen populären Gegenkandidaten aufgestellt. Sie sind verantwortliche Leiterin/verantwortlicher Leiter des Wahlteams „Eckerle“.

An die Beschreibung des Szenarios schließen sich die drei Subtest an.

Subtest 1: Situationsanalyse

Die Testpersonen erhalten eine kurze Beschreibung des Aufgabenkontextes sowie eine Anweisung zur Aufgabebearbeitung.

Hermann Eckerles Situation:

Für die morgige Podiumsdiskussion mit dem Herausforderer Dr. Igerle bittet Sie Ihr Chef Hermann Eckerle [...] anhand einer Reihe von Dokumenten einige seiner Statements zu überprüfen. Er möchte gerne wissen, ob seine Statements **so stimmen** und **so formuliert werden können**. Zudem sollen Sie angeben, **wie sicher Sie sich in Ihrem Urteil** sind. Hermann Eckerle bat Sie auch, dass Sie die Datenquellen angeben, die unbedingt notwendig für Ihr Urteil waren. [...] *Wichtig ist, Eckerle ist ein Mann mit klaren Prinzipien: Er möchte auch im Wahlkampf keine Aussagen machen, die nicht klar und eindeutig anhand der vorliegenden Informationen zu belegen sind.*

Nach der einführenden Beschreibung des Aufgabensettings und der Erläuterung der Aufgabebearbeitung stellt der Itemstamm die Aufgabe auf Itemebene dar. Der Subtest enthält insgesamt 9 Statements, die von der Testperson auf ihre Korrektheit hin überprüft werden müssen. Der Itemstamm wurde als Aussage formuliert, die anhand der im Begleitheft gegebenen Informationen auf ihre Richtigkeit hin überprüft werden muss. Anhand der Bearbeitung der Items zeigen die Testpersonen, ob sie die Fähigkeit besitzen, korrekte bzw. falsche oder unvollständige (und damit ebenfalls falsche) Schlüsse erkennen zu können. Damit kann festgestellt werden, ob die Testpersonen Informationen verknüpfen und anhand ihres Schlusses die jeweils gegebene Aussage überprüfen können. Die Aufgabebearbeitung zeigt, ob inkorrekte Kausalitäten erkannt sowie Übergeneralisierungen und dementsprechend inkorrektes induktives Schlussfolgern auf Basis unvollständiger Informationen vermieden werden kann. Darüber hinaus zeigen die beiden Itemsets zum Unterscheiden zwischen objektiven und subjektiven Informationen, ob die Testpersonen zwischen objektiven Fakten und subjektiven Meinungen unterscheiden und dementsprechend die Validität dieser Informationen erkennen können. Die Itemsets zum deduktiven Schlussfolgern zeigen, ob die Testpersonen mehrere Variablen deduktiv schließend miteinander verknüpfen können, um die gegebene Aussage zu überprüfen. Anhand der Items werden diejenigen kognitiven Kompetenzen überprüft, die im Rahmen der Analyse von Problemsituationen im fächerübergreifenden Kontext relevant werden können. Erst auf Basis der korrekten Analyse einer Ausgangssituation kann eine weitergehende Problembearbeitung erfolgen.

Die Testpersonen haben die Möglichkeit, die Items über eine vierstufige Skala zu beantworten (stimmt, ich bin mir sicher; stimmt, ich bin mir unsicher; stimmt nicht, ich bin mir sicher; stimmt

nicht, ich bin mir unsicher). Hintergrund dieser Konzeption war die Reduktion der Zufallsbeantwortung über die Darbietung von vier Antwortmöglichkeiten sowie die Berücksichtigung der Sicherheit der Testpersonen neben einer richtig-falsch Kategorisierung. Eine rationale Entscheidung beruht auf der Spezifizierung und Evaluation der gegebenen Informationen. Der Entscheider ist in der Lage, alle möglichen gegebenen Alternativen zu identifizieren, alle relevanten Zustände zu bestimmen, das Ergebnis jeder Kombination zu bestimmen, die einzelnen Ergebnisse zu bewerten, die Eintretenswahrscheinlichkeiten zu bestimmen und diejenige Entscheidung mit der maximalen Auftretenswahrscheinlichkeit zu wählen. Eine unsichere Entscheidung gibt Auskunft darüber, ob die Kompetenz, die Aufgabe zu lösen, auch in subjektiver Hinsicht gesichert ist. Dementsprechend zeigt die Beantwortung sowohl die Richtigkeit der Lösung, als auch die subjektiv empfundene Sicherheit der Entscheidung. Für die kompetenzbezogene Auswertung wurden die Antworten als richtig und falsch kategorisiert, ohne eine Berücksichtigung der Sicherheitskomponente. Eine Konfundierung mit psychologischen Variablen und eine Vermischung von objektiven und subjektiven Daten sollte ausgeschlossen werden.

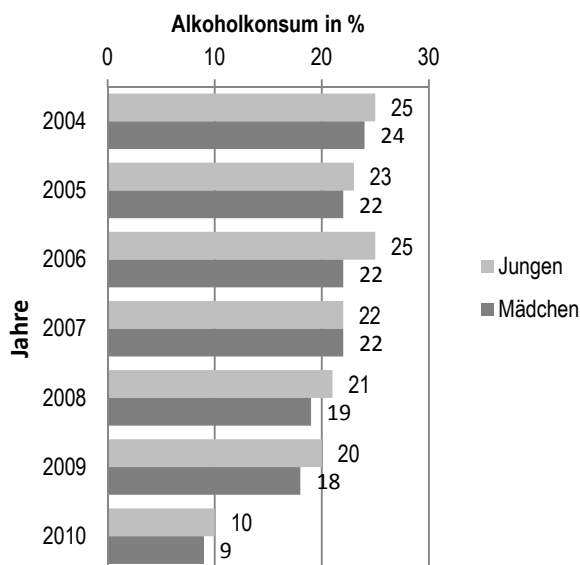
Insgesamt reicht zur Lösung der Aufgabe das einfache Erfassen der Informationen nicht aus. Vielmehr müssen die gegebenen Informationen korrekt verknüpft werden. Das Erkennen und Erfassen ist zwar eine notwendige, jedoch keine hinreichende Bedingung zur Lösung.

Das nachstehende Itembeispiel verdeutlicht die Aufgabenstruktur am Beispiel kausalen Schließens.

Itembeispiel

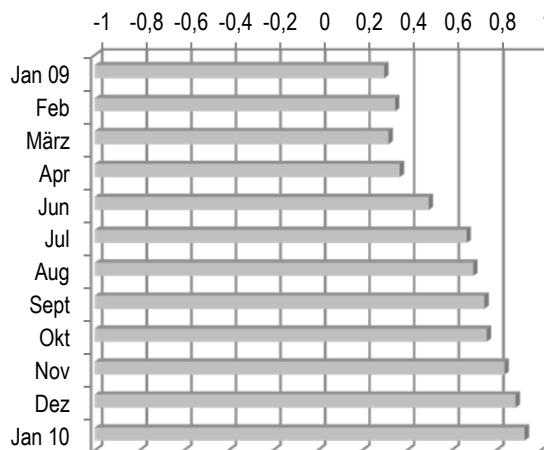
Verantwortlich für den Rückgang der Zahl betrunkenener Jugendlicher im Jahr 2010 sind die gestiegenen Alkoholpreise.

**Regelmäßiger Alkoholkonsum bei Jugendlichen
2004-2010, Beispielingen**



**Entwicklung der Alkoholpreise im Jahr 2009
gegenüber 2008 (in Prozent), Beispielingen**

Lesebeispiel: Alkoholpreise im Januar 2009 sind ca. 0,3% teurer als im Januar 2008



Antwortschema:

Hermann Eckerles Statements	Einordnung der Statements			
	Stimmt, kann so formuliert werden		stimmt nicht, kann so nicht formuliert werden	
	Ich bin mir sicher	Ich bin mir nicht sicher	Ich bin mir sicher	Ich bin mir nicht sicher
<p><i>Zur Kontrolle der Statements: Bitte Interne Unterlagen 1 prüfen</i></p>				
<p>Verantwortlich für den Rückgang der Zahl betrunkenen Jugendlicher im Jahr 2010 sind die gestiegenen Alkoholpreise.</p>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Subtest 2: Zielevaluation

Die Testpersonen erhalten auch hier eine kurze Beschreibung des Aufgabenkontextes sowie eine Anweisung zur Aufgabebearbeitung.

Hermann Eckerles Wahlkampfziele:

Hermann Eckerle bat Sie, dass Sie zwei der von ihm anvisierten Wahlkampfziele nochmals überprüfen. Er bittet Sie, dass Sie die nachstehenden Ziele dahingehend einschätzen, ob sie sowohl kostenneutral als auch zielführend für die Stadt sind. Zielführend bedeutet, dass das gesetzte Ziel erreichbar ist. Kostenneutral bedeutet, dass für die Stadt keine Kosten entstehen.

Alle Angaben die Sie benötigen, finden Sie in den **Internen Unterlagen 2**. Teilweise haben Sie sich selbst bereits einige Notizen zu den Zielen gemacht.

Nach dieser Einführung in das Aufgabensetting und nach der Erläuterung der Aufgabebearbeitung stellt auch hier der Itemstamm die Aufgabe auf Itemebene dar. Der Subtest enthält insgesamt 2 Items, die von der Testperson eingeschätzt werden müssen. Der Itemstamm wurde als Aussage formuliert. Anhand der Begleitinformation sowie unter Berücksichtigung der im Testheft gegebenen Notizen muss das Zielvorhaben im Hinblick auf die beiden Kriterien *Kostenneutralität* und *Zielführung* evaluiert werden. Die Bearbeitung der Items zeigen, ob die Testpersonen fähig sind, unter Berücksichtigung bestimmter Situationen Ziele anhand gegebener Kriterien einzuordnen und zu bewerten. Das heißt, diejenigen Informationen müssen erkannt werden, die notwendig sind, um das Erfüllen der Evaluationskriterien zu gewährleisten, einschränkende Bedingungen müssen berücksichtigt werden. Die gegebenen Informationen müssen verknüpft und verglichen werden. Über einen Abgleich mit den beiden Evaluationskriterien muss das zu erreichende Ziel letztlich bewertet werden. Die Testpersonen müssen aus vier gegebenen Antwortmöglichkeiten die richtige auswählen. Die Aufgabebearbeitung zeigt, ob Zielkonflikte erkannt werden. Es werden diejenigen kognitiven Kompetenzen überprüft, die für die Bewertung und Einordnung von Problemlösungszielen im fächerübergreifenden Kontext relevant sind.

Das nachstehende Itembeispiel verdeutlicht die Aufgabenstruktur und zeigt die zur Bearbeitung notwendigen Informationen, die im Begleitheft genannt werden.

Itembeispiel:

Die Lohnnebenkosten senken, um die Arbeitslosenzahlen bis 2015 zu senken.

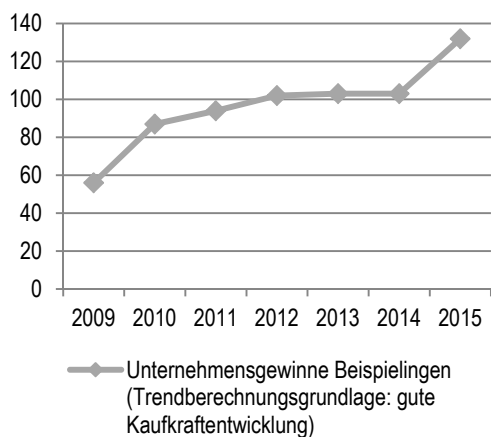
Sie haben folgende weitere Notizen:

Kostenneutral = städtische Investitionen bleiben auf dem bisherigen Stand

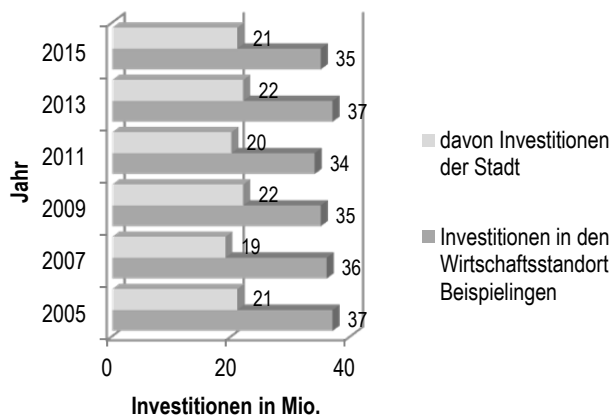
Die Lohnentwicklung wirkt sich direkt auf die Kaufkraft aus.

Gute Unternehmensbilanzen führen zu einem Ausbau der Stellen.

Gewinnentwicklung Beispieler Unternehmen
(Unternehmensbilanz: Hauptfaktor Gewinnentwicklung)



Investitionen in den Wirtschaftsstandort Beispieler (2005-2015)



Quelle: Agentur für Arbeit Beispieler



Statistik der Arbeitsagentur Beispieler

Dokument Nr. 23az.4		
Agentur	Arbeitslose	Prozentangabe
Agenturbezirk gesamt	18.178	5,3
Hauptagentur Beispieler	14.077	6,4
Agentur Teilbeispieler	2.145	3,0
Agentur Partbeispieler	1.956	3,9
Stand: September 2010		

Dokument Nr. 23az.5b			
Veränderungen gegenüber dem vergangenen Jahresquartal	September 2010	Veränderungen gegenüber Vorjahr	Prognose 2015
- 0,4	5,3% Arbeitslosenquote	- 0,4	- 0,1

- 1.1856	18.187 Arbeitslose	- 1.086	k. A.
- 283	1.828 jüngere Arbeitslose unter 25 Jahre	- 518	k.A.
- 90	5.341 ältere Arbeitslose über 50 Jahre	+ 170	k.A.
- 660	5.958 Abmeldungen aus Arbeitslosigkeit	+ 988	k.A.
+ 372	3.359 offene Stellen im Bestand	+ 199	k.A.
+ 385	2.057 Stellenzugänge	+ 272	k.A.
Trendberechnung: Voraussetzung positive Stellenentwicklung			

Trendbericht Unternehmensvertretung
 Das Lohnniveau wird in den nächsten Jahren sinken. Entlastend könnten zwar geringere Lohnnebenkosten wirken, die Stadt hat dies jedoch bislang noch nicht in Aussicht gestellt.

Antwortschema:

Kreuzen Sie an...

Das Vorhaben ist	zielführend	nicht zielführend
Kostenneutral		
nicht kostenneutral		

Subtest 3: Handlungsplanung

Die Testpersonen erhalten äquivalent zu den bisherigen Aufgaben einleitend eine Beschreibung des Kontextes sowie eine Anweisung zur Aufgabebearbeitung:

Hermann Eckerles nächste Schritte:

Sie möchten wissen, wie Sie Ihr Wahlkampfteam so gut wie möglich auf die nächsten Wahlkampfschritte vorbereiten können. Zu den verschiedenen Themen [...] haben Sie sich die wichtigsten Schritte bereits notiert. Bevor Sie Ihrem Team die nächsten Schritte nennen können, müssen Sie Ihre Stichpunkte nochmals durchgehen und die geplanten Maßnahmen in die richtige Reihenfolge bringen [...]. Schritte, die für die Maßnahme nicht notwendig sind, sind für Eckerle im Wahlkampf nicht notwendig und sollen gestrichen werden.

Beispiel: Die richtige Reihenfolge der Maßnahmen könnte sein: X-Y-Z; nicht notwendig ist Schritt Z, deshalb kann Schritt Z gestrichen werden: ~~Z~~

Sollten Sie Korrekturen an Ihren Antworten vornehmen, streichen Sie bitte klar durch und notieren den neuen Buchstaben darunter oder darüber: B

~~X~~

Nach der Einführung in das Aufgabensetting und der Erläuterung der Aufgabebearbeitung stellt der jeweilige Itemstamm die Aufgabe auf Itemebene dar. Der Subtest enthält insgesamt 2 Items, die von der Testperson bearbeitet werden müssen. Der Itemstamm wurde als Ziel formuliert. Anhand der Begleitinformation, die im Testheft notiert sind muss die gegebene Information, das heißt die nächsten Wahlkampfschritte, so gereiht werden, dass das Ziel, das im Itemstamm genannt wurde, erreicht werden kann. Die Bearbeitung der Items zeigt, ob die Testpersonen unter Berücksichtigung der gegebenen Situation bestimmte Handlungen planen kann, das bedeutet, bestimmte Gegebenheiten so reihen kann, dass ein vorgegebenen Ziel konfliktfrei erreicht werden kann. Hierzu muss eine komplexe Situation in Einzelschritte zerlegt werden. Nach dieser Zerlegung müssen die gegebenen Informationen verknüpft und in diejenige Reihenfolge gebracht werden, die eine Zielerreichung erlaubt. Dabei müssen einschränkende Bedingungen erkannt und berücksichtigt, überflüssige Informationen erkannt und gestrichen werden.

Anhand dieser Testaufgaben werden kognitive Kompetenzen überprüft, die für die Analyse und Zerlegung komplexer Situationen relevant sind und die es erlauben, Problemlöseschritte im fächerübergreifenden Kontext zu planen. Da dieses Aufgabenset schwieriger als Subtest 1 und 2 ist, wurde es an das Testende gesetzt.

Insgesamt erfordert die richtige Beantwortung mehr als das Erkennen der relevanten Information: Erst ein Vergleich der gegebenen Informationen lässt das Erkennen der relevanten Information zu

und erlaubt einen Ausschluss der irrelevanten Information sowie eine korrekte Platzierung der Handlungsschritte. Auch Distraktoren müssen erkannt werden: Distraktoren enthalten Informationen, die nicht mit den im Text genannten Informationen übereinstimmen (proximale Distraktoren) oder auch im Text nicht genannt werden (distale Distraktoren). Die Distraktoren beziehen sich auf Informationen, die für die Handlungsplanung und damit die Zielerreichung irrelevant sind. Das nachstehende Itembeispiel verdeutlicht die Aufgabenstruktur.

Itembeispiel:
Die Arbeitslosigkeit weiter senken.

Ihre Stichpunkte zum Thema:

- Unternehmen investieren in den Stellenausbau, wenn ihre Unternehmensbilanz positiv ausfällt.
- Voraussetzung dafür ist, dass die Stadt sinkende Lohnnebenkosten signalisiert.
- Ein Stellenausbau durch Unternehmen ist unabdingbar für sinkende Arbeitslosenzahlen.
- Die städtischen Finanzen (das *Stadtsäckel*) erlauben es, entweder die Abgaben für Wirtschaftsunternehmen (hier: Lohnnebenkosten) zu senken oder in die Bildung zu investieren.
- Die Bürger akzeptieren sinkende Lohnnebenkosten nur dann, wenn die Stadt vorher verspricht, in die Bildung zu investieren.
- Die Stadt kann Werbemaßnahmen genehmigen, um ihr Budget wieder für eine weitere Initiative aufzufüllen. Um das städtische Image zu wahren, muss das Stadtsäckel vor Genehmigung leer sein.
- Es dürfen nicht zwei Werbemaßnahmen hintereinander durchgeführt werden.

Folgende Maßnahmen sind geplant. Bringen Sie diese in die richtige Reihenfolge. Streichen Sie ggf. irrelevante Maßnahmen (Bsp. X)	
Signal an Unternehmen: Löhne sollen gesenkt werden	A
Packet 1: Werbemaßnahmen der Stadt	B
Packet 2: Werbemaßnahmen der Stadt	C
Bildungsinvestitionen tätigen	D
Senkung der Lohnnebenkosten an Unternehmen signalisieren	E

Antwortschema:

Die Maßnahmen sollen in folgender Reihenfolge durchgeführt werden, sodass das Ziel die *Arbeitslosigkeit senken* erfolgreich erreicht werden kann (Bsp. X – Y – Z):

Testmodell

Über das Aufgabenspektrum wurde versucht, fächerübergreifende Problemlösekompetenz im Sinne der Erfassung relevanter kognitiver Teilkompetenzen zentraler Problemlösephasen zu messen. Die Beschreibung der Subtests und Itemsets stellte die jeweils operationalisierten Kompetenzen dar. Dabei zeigt sich, dass im Hinblick auf grundlegende Fähigkeiten wie das Erkennen und Verknüpfen von Informationen, bzw. das Erkennen spezifischer Einschränkungen von einer Interkorrelation der Subtests ausgegangen werden kann. Deshalb wurden in den Subtests nicht nur notwendige kognitive Kompetenzen, sondern für die jeweils operationalisierte Problemlösephase vor allem hinreichende kognitive Kompetenzen fokussiert. Das nachstehende Modell fasst nochmals die Struktur fächerübergreifender Problemlösekompetenz wie sie im Rahmen des Testverfahrens abgebildet wurde, zusammen. Mit Blick auf die Analyse unterschiedlicher Kompetenzen wurde versucht, unterschiedliche Schwierigkeiten abzubilden. Die nachstehende tabellarische Übersicht zeigt zusammenfassend über welche Kriterien die Schwierigkeit jeweils variiert wurde.

Grafik 5: Erfasste Teilkompetenzen beim fächerübergreifenden Problemlösen

Grafik 6: Übersicht über alle Items des Testverfahrens

Subtest 1	Situationsanalyse	Induktives Schlussfolgern	Handwerker
			Migrationsbiographie
		Deduktives Schlussfolgern	Berufstätige Frauen
			Familienunternehmen
			Gesundheitskosten
		Kausalität und Korrelation	Verkehrsunfälle
			Alkoholkonsum
		Fakten und Meinung	Arbeitslosenzahlen
Jugendliche Raucher			
Subtest 2	Zielevaluation	Bewertung und Einschätzung von Zielen	Zufriedenheit der Bürger
			Lohnnebenkosten
Subtest 3	Handlungsplanung	Reihung und Platzierung von Handlungsschritten	Arbeitslosigkeit
			Getränkehersteller

9.3 Ergebnisraum und Auswertungsschema

Für Subtest 1 wurde für jede richtige Antwort ein Punkt vergeben. Es konnten insgesamt 9 Punkte erreicht werden. Für Subtest 2 wurde je nach richtiger Evaluation des Ziels ein Punkt vergeben. Pro Item musste das vorgegebene Ziel anhand von zwei Kategorien evaluiert werden. Es konnten pro Aufgabe 2 Punkte, insgesamt für Subtest 2 vier Punkte erworben werden. Bei der Auswertung der Umordnungsaufgaben in Subtest 3 wurde die richtige Platzierung eines jeden Schrittes als richtige Antwort gewertet. Jeder richtig erkannte Distraktor wurde ebenfalls als richtige Antwort gewertet. Insgesamt konnten 5 Punkte pro Item, 10 Punkte insgesamt erreicht werden.

9.4 Umgang mit fehlenden Werten

Die Gründe für fehlende Werte können unterschiedlich sein (Bortz, 2005): So können Werte fehlen, da nicht alle Items allen Testpersonen vorgelegt wurden, fehlende Werte können entstehen, da Items aus zeitlichen Gründen nicht mehr bearbeitet werden konnten und von den Testpersonen absichtlich ausgelassen wurden oder auch, da bestimmte Items als zu schwierig empfunden wurden. Fehlende Werte liegen dementsprechend nicht nur im Datenverlust oder dem Auslassen von Ausreißern begründet, sondern auch im Instrument, in der Testleitung oder auch der Testperson. Die Problematik im Zusammenhang mit dem Umgang mit fehlenden Werten und ihrer Ersetzung über *missing-data* Techniken orientiert sich deshalb daran, ob die fehlenden Werte systematisch gehäuft auftreten oder als zufällig betrachtet werden könnten (Bortz, 2005). Um möglichen Verzerrungen vorzubeugen, müssen deshalb nicht nur Gründe für fehlende Werte

diskutiert werden, sondern auch angemessene Ersetzungsmethoden. Neben klassischen Methoden des listen- oder paarweisen Ausschlusses werden Imputationsmethoden vorgeschlagen. Baker (2004) nennt die Möglichkeit, fehlende Werte unter bestimmten Umständen als falsche Antworten zu definieren. Dieser Annahme, die lediglich für Leistungstests gilt, liegt zugrunde, dass Testpersonen die entsprechenden Aufgaben nicht versehentlich oder aus Motivationsgründen ausließen, sondern sie nicht beantworten konnten. Dies setzt voraus, dass die Testpersonen in der Instruktion aufgefordert wurden, alle Items zu beantworten und am Ende der Testung nochmals prüfen sollten, ob dies auch der Fall ist. Im Rahmen der Hauptstudie wurden die hier vorliegenden fehlenden Werte analysiert. In diesem Zusammenhang wird auf das gewählte Ersetzungsverfahren nochmals eingegangen.

10. Pilotierung des Testinstruments

Um von den manifesten Äußerungen der Testpersonen zuverlässig auf das latente Merkmal der Ausprägung fächerübergreifender Problemlösekompetenz schließen zu können, muss die Güte des Testverfahrens überprüft werden. Im Folgenden wird dargestellt, inwieweit das vorliegende Instrument geeignet ist, fächerübergreifende Problemlösekompetenz zu messen. Dabei wird auf die Kennwerte der Vorstudie sowie der Hauptstudie Bezug genommen. Einleitend wird die der Vorstudie zugrundeliegende Stichprobe und die Durchführung der Pilotierung dargestellt. Im Zentrum der Vorstudie stand die Auswahl geeigneter, unterschiedlich schwieriger Items für die Hauptstudie. Das Testverfahren musste insgesamt gekürzt werden.

10.1 Stichprobe der Vorstudie

Für die Vorstudie konnten insgesamt 54 Studierende gewonnen werden. Um die Schwierigkeit der Items einordnen zu können, richtete sich die Vorstudie an Studierende zwischen dem 5. und 7. Semester, das heißt, an Studierende kurz vor dem Abschluss. Zum Zeitpunkt der Durchführung der Pilotierung (Sommersemester) wurden Studierende des fünften Bachelor- und siebten Lehramtsemesters befragt. Um eine Konfundierung mit Studierenden auszuschließen, die an der Hauptstudie teilnehmen sollten, wurden Studierende aus Fachbereichen der Kath. Universität Eichstätt-Ingolstadt rekrutiert, die nicht Teil der Hauptstudie sein würden. Aus dem fünften Semester beteiligten sich $N=35$ Studierende aus den Studiengängen der Betriebswirtschaftslehre ($N=17$), der Geographie ($N=8$) und den Politikwissenschaften ($N=10$). Des Weiteren wurden $N=19$ Studierende des siebten Semesters des Studiengangs der Schulpsychologie rekrutiert. Insgesamt nahmen 32 weibliche und 20 männliche Testpersonen teil, für zwei Studierende fehlten entsprechende Angaben. Das Durchschnittsalter der Studierenden lag bei 23,24 Jahren ($SD= 2,16$), 5 Studierende gaben ihr Alter nicht an.

10.2 Durchführung der Vorstudie

Die Vorstudie fand im Rahmen der zur Verfügung gestellten Lehrveranstaltungszeit statt und umfasste insgesamt 60 Minuten. Die Studierenden wurden über das Ziel der Studie aufgeklärt und bezüglich der Aufgabenbearbeitung instruiert. Den Testunterlagen vorangestellt war diejenige Instruktion, die auch der Hauptstudie zugrundelag. Der formale Testaufbau, das inhaltliche Setting sowie das Testlayout entsprach dem der Hauptstudie. Aufgrund der umfangreicheren Itemzahl für die Vorstudie und dem größeren Testumfang konnten die Studierenden im Rahmen der Pilotierung nicht das gesamte Testverfahren, das heißt alle drei Subtests, bearbeiten, die Bearbeitungszeit hätte mehr als 90 Minuten in Anspruch genommen. Für die Pilotierung wurde deshalb drei gleich großen Studierendengruppen je

ein Subtest zur Bearbeitung vorgelegt. Jeder Subtest nahm trotz unterschiedlicher Itemanzahl in etwa die gleiche Bearbeitungszeit, zwischen 30 und 45 Minuten, in Anspruch.

Das Testverfahren bestand in der Vorstudie aus insgesamt 27 Items. Subtest 1 umfasste 19 Items, wobei jeweils vier Items deduktivem Denken, subjektiver und objektiver Informationen unterscheiden sowie Korrelation und Kausalität unterscheiden zuzuordnen sind. Ein Itemset von jeweils drei Items lag zum induktiven sowie abduktiven Denken vor. Das Itemset zum abduktiven Denken wurde als Sonderform des induktiven Denkens in der Hauptstudie nicht mehr aufgenommen. Für die Subtests 2 und 3 lagen Itemsets von jeweils vier Items vor.

Subtest 1 (Situationsanalyse) bearbeiteten insgesamt $N=20$ Studierende, Subtest 2 (Zielevaluation) bearbeiteten ebenfalls $N=20$ Studierende, Subtest 3 (Handlungsplanung) wurde von $N=16$ Studierenden bearbeitet. Intention der Vorstudie war die Überprüfung der Item- und Skalenkennwerte, sowie die Frage, welche Items trotz der Kürze das Testverfahren eine möglichst differenzierte Aussage hinsichtlich Schwierigkeiten beim fächerübergreifenden Problemlösen ermöglichen. Die Begrenzung der Testzeit über die Reduktion der Itemanzahl hat für die Hauptstudie eine Verschlechterung der Reliabilität zur Folge, die jedoch in Kauf genommen wurde. Aufgrund der bereits erläuterten Dreiteilung des Testverfahrens konnte im Rahmen der Vorstudie keine Gesamtreliabilität berechnet oder ein Globalindex zur fächerübergreifenden Problemlösekompetenz gebildet werden. Ebenso wenig konnten strukturprüfende Verfahren für den Gesamttest erfolgen.

10.3 Ergebnisse der Vorstudie

Im Folgenden wird begründet, welche Items für die Hauptstudie ausgewählt wurden und inwieweit sie ggf. verändert wurden. Dabei wird Bezug genommen auf Itemschwierigkeiten, Item-Skalen-Korrelationen, Itemtrennschärfen sowie Itemvarianzen. Zu berücksichtigen ist, dass nicht eindeutig beurteilt werden kann, inwieweit sich die Itemkennwerte von Vorstudie zu Hauptstudie verbessert haben, da einerseits die Items für den Haupttest teilweise überarbeitet wurden und es sich andererseits zwar um ähnlich zusammengesetzte, aber unterschiedliche Stichproben in Vor- und Hauptstudie handelt.

In erster Linie sollten die Schwierigkeitsindices überprüft werden: Der Schwierigkeitsindex P_i eines Items i ist der Quotient aus der bei diesem Item tatsächlich erreichten Punktsomme aller n Probanden und der maximal erreichbaren Punktsomme aller n Probanden bei diesem Item multipliziert mit 100 (Kelava & Moosbrugger, 2007). Itemschwierigkeiten um 50 gelten als optimal (Kelava & Moosbrugger, 2007). Niveautests sollten Items unterschiedlicher Schwierigkeiten sowie ggf. auch Extreme umfassen. Zur Beurteilung der Itemschwierigkeiten und Auswahl geeigneter Items wurde für die vorliegende Studie das Kriterium angelegt, dass die Itemschwierigkeit der Fähigkeit der anvisierten Stichprobe angepasst sein sollte, wobei unterschiedliche Kompetenzausprägungen erfasst

werden sollten. Weiterhin sollten möglichst dichotome Items ausgewählt werden, um für die Analyse fächerübergreifender Problemlösekompetenzen bei Studierenden sowohl Stärken als auch Schwierigkeiten explorieren zu können.

In Zusammenhang mit der Schwierigkeit der Items steht allerdings die Varianz der Items. Testaufgaben extremerer Schwierigkeit ermöglichen weniger Differenzierung als Testaufgaben mit mittlerer Schwierigkeit. Daneben drückt auch die Trennschärfe die Aussagekraft des Items aus. Ein Item mit hoher Trennschärfe differenziert gut zwischen Personen unterschiedlicher Eigenschaftsausprägungen, Items mit geringer Trennschärfe können dies weniger (Rost, 2004). Die klassische Trennschärfe ist definiert als die Korrelation eines Items i mit dem Testergebnis t (Rost, 2004a). Trennschärfen sollten zwischen .4 und .7 liegen und .3 möglichst nicht unterschreiten (Kelava & Moosbrugger, 2007). Zur Beurteilung der Trennschärfe wurden mehrere Aspekte berücksichtigt: Das Konzept der Trennschärfe als Indikator der prognostischen Validität eines einzelnen Items beruht in erster Linie auf einem relativ homogen konstruierten Verfahren (Bortz & Döring, 2005). Zudem ist die Stabilität eines Trennschärfenkoeffizienten erst ab größeren Stichproben einigermaßen gesichert. Bei einem heterogen konstruierten Verfahren wie im vorliegenden Fall ist demzufolge die Trennschärfe relativ labil. Zudem ist die Probandenzahl in der vorliegenden Vorstudie gering. Die Trennschärfe- und Reliabilitätskoeffizienten müssen vor diesem Hintergrund interpretiert werden.

Die Tabellen im Anhang geben einen Überblick über die Kennwerte der Subtests (Anhang, Tabelle 1-10). Dabei ist angegeben, welche Items für die Hauptstichprobe ausgewählt wurden und inwieweit sich die Kennwerte im Rahmen der Hauptstudie verändert haben. Im Folgenden wird die Auswahl der Items begründet. Die Tabellen im Anhang zeigen für die ausgewählten Items auch die Kennwerte der Hauptstichprobe. Die Itemantworten sind in Vorstudie und Hauptstudie für Subtest 1 binär kodiert (falsch = 0, richtig = 1), für Subtest 2 wird für jede richtige Antwort 1 Punkt vergeben (falsch = 0; 1 richtig Antwort = 1; 2 richtige Antworten = 2). Für Subtest 3 wird für jede richtige Platzierung 1 Punkt vergeben (keine richtige Platzierung = 0; 1 richtige Platzierung = 1; ... 5 richtige Platzierungen = 5), zudem wurde das Erkennen der Distraktoren ausgewertet.

Kennwerte des Subtests Situationsanalyse

Im Rahmen der Vorstudie weist Subtest 1 (Anhang, Tabelle 1-3) insgesamt eine gute Reliabilität von Cronbachs Alpha .74 auf, die Schwierigkeit der Items liegt zwischen schwierig ($P_1 = 25,0$, Item *Tourismus*, Induktion) und sehr leicht ($P_1 = 90,0$, Item *Nichtraucher*, Fakten vs. Meinung), wobei die überwiegende Anzahl der Items eine mittlere Schwierigkeit aufweist.

Das Itemset zum deduktiven Schlussfolgern sollte Items mit aufsteigender Schwierigkeit beinhalten, deshalb wurden drei Items ausgewählt, um eine größere Bandbreite an Schwierigkeitsindices beim deduktiven Schließen zu evaluieren. Die Vorstudie zeigt, dass die Items – bis auf das Item *Gesundheit*

– durchschnittlich schwierig sind und entsprechend differenzieren. Aus Konstruktionsgründen wurden die Testaufgaben *Wahlkampf* sowie *Bewerbungskurse* für die Hauptstudie ausgeschlossen: Die Testaufgabe *Wahlkampf* wurde in der Vorstudie aufgenommen, um analoges Problemlösen zu simulieren. Um die Kohärenz der Items zum deduktiven Schließen zu gewährleisten wurde das Item gestrichen. Die Testaufgabe *Berufstätigkeit* wurde dem Item *Bewerbungskurse* aufgrund inhaltlicher Kohärenz vorgezogen, die Items *Gesundheit* und *Familienunternehmen* wurden trotz schwächerer Trennschärfe aufgrund ihrer höheren Schwierigkeit ausgewählt und nochmals überarbeitet, um die Kennwerte zu verbessern: Für die Hauptstudie mussten zur Lösung des Items *Gesundheit* Informationen aus zwei Graphiken und einer Tabelle entnommen und verknüpft werden. Auf dieser Basis musste evaluiert werden, ob die Aussage im Aufgabenstamm richtig ist. Das Item *Gesundheit* wurde als schwierigstes Item konstruiert. Die Testaufgabe *Berufstätigkeit* wurde vereinfacht, sodass nur noch zwei Graphiken ausgewertet und verknüpft werden mussten. Die Aussage des Itemstamms, die eine Einschränkung beinhaltet, musste mit den Ergebnissen der Auswertung in Verbindung gesetzt und überprüft werden. Die Testaufgabe *Familienunternehmen* sollte die leichteste der Testaufgaben zum deduktiven Schlussfolgern darstellen. Hier wurde die Aufgabe so vereinfacht, dass zwei Graphiken ausgewertet und in Beziehung gesetzt werden mussten, um die im Itemstamm gegebene Aussage ohne weitere einschränkende Bedingung zu überprüfen. Tabelle 2 im Anhang zeigt, dass im Rahmen der Hauptstudie die intendierte aufsteigende Schwierigkeit abgebildet werden konnte, das Item *Gesundheit* jedoch weiterhin sehr schwierig zu lösen war. Die Darstellung der Befunde der Studie wird auf die Schwierigkeiten bei der Lösung der Items eingehen. Mit Blick auf die differierenden Kennwerte der Items bei einem Vergleich der Vor- und Hauptstudie ist auch danach zu fragen, warum die Studierenden der Wirtschaftswissenschaften (Vorstudie) etwas weniger Probleme bei der Bearbeitung dieser Items hatten als die Studierenden der Sozial- und Geisteswissenschaften (Hauptstudie). Eine Interferenz des Studiengangs kann nicht ausgeschlossen werden, vor allem auch, da im Rahmen der Hauptstudie Studierende der Sozial- und Geisteswissenschaften vereinzelt rückmeldeten, dass die Aufgaben zum deduktiven Schließen dem wirtschaftlichen Bereich entlehnt empfunden wurden und deshalb als sehr schwierig empfunden worden seien. Die Rückmeldungen wurden aufgrund der sehr geringen Anzahl nicht dokumentiert.

Für das Itemset zur Unterscheidung zwischen subjektiven und objektiven Informationen wurden die Items *Raucher* und *Arbeitslose* ausgewählt: Das Item *Gentechnik* war in einen emotional gefärbten Kontext eingebettet, das Item *Jugendliche Wähler* enthielt zusätzlich die Anforderung des deduktiven Schließens, beide Items wurden deshalb für die Hauptstudie ausgeschlossen. Der Einbezug weiterer Schwierigkeitsmomente wie bspw. emotional gefärbter Inhalte wäre interessant, wurde für die Hauptstudie jedoch ausgeschlossen, um ausschließlich die Fähigkeit zur Unterscheidung zwischen subjektiver und objektiver Information zu messen. Die Items *Raucher* und *Arbeitslose* wurden für die Hauptstudie nicht mehr verändert: Die im Itemstamm gegebene Aussage musste über einen Vergleich einer subjektiven und einer objektiven Informationsquelle überprüft werden. Die subjektive

Information war in beiden Fällen eine Aussage des Bürgermeisters. Die objektive Information stellte in beiden Fällen eine Statistik des fiktiven Statistikamtes dar. Die Schwierigkeit der beiden Aufgaben wurde unterschiedlich gestaltet. Die objektive Information bei der Testaufgabe *Raucher* stützt im Gegensatz zur Testaufgabe *Arbeitslosigkeit* auf den ersten Blick die subjektive Meinung (Sinken des Nichtraucheranteils). Diese Schwierigkeit spiegelt sich im Schwierigkeitsindex wider, auch in der Hauptstudie ist das Item deutlich schwieriger. Die Trennschärfeindices verbessern sich allerdings deutlich.

Für die Itemsets zum induktiven Denken sowie zum Unterscheiden zwischen Korrelation und Kausalität wurden die Items *Alkohol*, *Neuwagen* (Korrelation vs. Kausalität) sowie *Handwerker* und *Migration* (Induktives Schlussfolgern) ausgewählt. Bei den Items zum Unterscheiden zwischen Korrelation und Kausalität sollte gewährleistet sein, dass sie eindeutig zu beantworten sind. Die Itemsets wurden für die Hauptstudie nicht mehr überarbeitet: Zur Überprüfung der im Itemstamm gegebenen Aussage lagen bei den Testaufgaben zum korrelativen vs. kausalen Denken zwei Statistiken vor, die die im Itemstamm gegebenen Informationen zwar bestätigten, die jedoch einem kausalen Zusammenhang entsprechen und weitere Erklärungsbedingungen vernachlässigen würden. Zur Bearbeitung des Itemsets zum induktiven Denken lag jeweils eine Information zur Überprüfung der Aussage im Itemstamm vor. Eine Bestätigung der im Itemstamm gegebenen Aussage würde einer Übergeneralisierung der Einzelinformation entsprechen.

Bei korrelativen vs. kausalen Schließen wurde aufgrund ungenügender Trennschärfe das Item *Drogen* eliminiert, das Item *Politiker* ebenso, zugunsten der Beibehaltung der Items *Alkohol* und *Neuwagen*. Über diese beiden Items konnte nicht nur die Frage des Erkennens korrelativer Zusammenhänge eruiert werden. Vielmehr ist die Intention, zu überprüfen, inwieweit die Plausibilität des Zusammenhangs eine Rolle spielt: Der Zusammenhang zwischen Alkoholkonsum und Preiserhöhungen könnte, so die Annahme, plausibler erscheinen als der Zusammenhang zwischen Neuwagenverkäufen und Unfallzahlen. Bei den Items zum induktiven Schließen wurde aufgrund eines deduktiven Anteils das Item *Tourismus* für die Hauptstudie ausgeschlossen. Die Übersicht über die Kennwerte der Hauptstudie bestätigt die Auswahl: Schwierigkeit und Differenzierungsfähigkeit sind angemessen. Die Items erscheinen in der Schwierigkeit jedoch homogen.

Tabelle 3 im Anhang zeigt Schwierigkeits- und Trennschärfeindices für die zu Itemsets zusammengefassten Items. Hier bildet sich nochmals die Heterogenität des Itemsets zum deduktiven Schließen ab. Die Reliabilität der Gesamtskala ist aufgrund dessen schwach und verbessert sich mit Ausschluss des Itemsets zum deduktiven Schließen entsprechend.

Kennwerte des Subtests Zielevaluation

Subtest 2 Zielevaluation (Anhang, Tabelle 4-5) weist mit Cronbachs Alpha .39 eine geringe Reliabilität auf. Die Schwierigkeit der Items ist angemessen, das Item *Lohnnebenkosten* ist mit $P_1=$

25,0 das schwierigste. Die Items können mit Trennschärfen zwischen $r_{it}=6$ und $r_{it}=7$ gut differenzieren. Für die Hauptstudie ausgewählt wurden jeweils das schwierigste sowie eines der einfacheren Items. Dabei wurde nochmals geprüft, welches der Items am besten einer Simulation beruflicher Anforderungen im Sinne der Strukturierung komplexer Information entspricht. Vor diesem Hintergrund wurden die Items *Lohnnebenkosten* und *Verwaltungsapparat* für die Hauptstudie ausgewählt. Die ausgewählten Items wurden für die Hauptstudie nochmals vereinfacht, wobei das Item *Verwaltungsapparat* als das einfachere der beiden Testaufgaben konzipiert wurde. Zur Bearbeitung der Testaufgabe *Verwaltungsapparat* müssen in der Hauptstudie drei Informationen ausgewertet und verknüpft sowie mit dem gegebenen Verwaltungsorganigramm in Verbindung gesetzt werden. Das Ergebnis dieses Schlusses dient dazu, unter Abgleich der beiden definierten Evaluationskriterien *kostenneutral* und *zielführend*, die im Itemstamm gegebene Aussage zu evaluieren. Zur Bearbeitung der Testaufgabe *Lohnnebenkosten* liegen vier Informationen vor, die ausgewertet und verknüpft werden müssen. Dabei geben zwei weitere Informationen spezifische Einschränkungen und damit die Art der Verknüpfung der Informationen vor. Auch hier muss das im Itemstamm genannte Ziel über einen Abgleich des Ergebnisses mit den beiden Evaluationskriterien *kostenneutral* und *zielführend* eingeordnet werden. Die Kennwerte der Hauptstudie in Tabelle 5 zeigen für das Itemset eine mittlere Schwierigkeit und gute Trennschärfe, wobei beide Items gleich schwierig sind.

Kennwerte des Subtests Handlungsplanung

Subtest 3 zeigte im Rahmen der Vorstudie eine hohe Varianz, gute Trennschärfe und insgesamt eine gute Reliabilität mit Cronbachs Alpha $=.71$ (Anhang, Tabellen 6-10). Die Items sind jedoch sehr schwierig mit Kennwerten zwischen $P_1=12,5$ und $P_1=33,3$. Die Schwierigkeit war ausschlaggebend dafür, die Items nochmals zu überarbeiten. So wurde die Anzahl der zu manipulierenden Informationen, die Anzahl der Distraktoren sowie die Anzahl der zu platzierenden Maßnahmen reduziert. Aufgrund der besseren inhaltlichen Passung zu beruflich relevanten fächerübergreifenden Problemlösekompetenzen wurden für die Hauptstudie die Items *Arbeitslosigkeit senken* und *Getränkehersteller* ausgewählt und folgendermaßen überarbeitet: Für beide Items wurden die Informationen zur Situationsbeschreibung auf sechs gekürzt. Diese Informationen müssen verknüpft werden. Gegeben wurden für jedes Item fünf Maßnahmen A bis E, darin waren jeweils zwei Distraktoren enthalten. Der Itemstamm jedes Items benennt ein Ziel, das vor dem Hintergrund der jeweils beschriebenen Situation durch die korrekte Platzierung der Maßnahmen A bis E erreicht werden muss. Das heißt, die zur Zielerreichung notwendigen Maßnahmen A bis E müssen ausgewählt und in die richtige Reihenfolge gebracht werden. Die enthaltenen zwei Distraktoren, also die nicht notwendigen Maßnahmen, müssen erkannt und gestrichen werden. Die Schwierigkeitsanalyse für die Hauptstudie zeigt, dass auch nach der Vereinfachung der Aufgabe die Schwierigkeit nicht verringert

werden konnte. Äquivalent zu den Items des deduktiven Schließens stellt sich die Frage, inwieweit Unterschiede zwischen Wirtschaftswissenschaftsstudierenden und Sozialwissenschafts- sowie Geisteswissenschaftsstudierenden unterschiedliche Schwierigkeitsindices in Pilotierung und Hauptstudie nach sich ziehen. Eine detailliertere Auswertung im Rahmen der Befunddarstellung zeigt, inwieweit Schwierigkeiten näher beschrieben werden können. Auffallend ist, dass sich unter Miteinberechnung des Erkennens der beiden Distraktoren die Schwierigkeit nochmals erhöht (Anhang, Tabelle 8). Tabellen 9 und 10 zeigen, inwieweit die Distraktoren im Rahmen der Hauptstudie erkannt wurden. Proximaler Distraktor ist eine Maßnahme, die in engem inhaltlichen Zusammenhang zur Aufgabe steht, jedoch unnötig ist. Distaler Distraktor, wenn sie in weniger direktem inhaltlichen Zusammenhang steht. Tabellen 9 und 10 zeigen, dass bei beiden Items der proximale Distraktor eher als der distale Distraktor erkannt wird. Das heißt, eine in engerem inhaltlichen Zusammenhang stehende Maßnahme wurde eher als irrelevant erkannt. Dies lässt die Annahme zu, dass Maßnahmen eher beibehalten wurden, wenn ihre inhaltliche Notwendigkeit nicht direkt ausgeschlossen werden konnte. Im Hinblick auf eine Ungleichverteilung zeigen sich keine auffälligen Häufungen.

Abschließend kann berichtet werden, dass die Schwierigkeit des Itemsets dann einen akzeptablen Wert von $P_i=54,57$ (*Arbeitslosigkeit senken*) bzw. von $P_i=46,45$ (*Getränkehersteller*) erreicht, wenn zwischen Nichtlösern und Lösern, die mindestens eine richtige Platzierung vornahmen, differenziert wird. Über die Items zur Strukturierung komplexer Informationen (Handlungsplanung) können damit insgesamt unterschiedliche Differenzierungen vorgenommen werden: Einerseits sowohl zwischen Lösern und Nichtlösern, andererseits zwischen Lösern mit unterschiedlicher Anzahl von korrekten Platzierungen. Darüber hinaus kann auch über die Distraktoren die Schwierigkeit bestimmt werden. Die besondere Schwierigkeit der Aufgabe, so kann jedoch festgehalten werden, liegt in der vollständig korrekten Platzierung der Maßnahmen unter Mitberücksichtigung des Erkennens der Distraktoren.

10.4 Umgang mit fehlenden Werten

Tabelle 12 im Anhang zeigt eine Übersicht der fehlenden Werte in der Hauptstudie auf Ebene der einzelnen Items, Tabelle 11 auf Ebene der Itemsets. Es ist feststellbar, dass eine Systematik fehlender Werte, das heißt, eine Häufung über 5% (Lüdtke, Robitzsch, Trautwein & Köller, 2007) bei den Items *Lohnnebenkosten* aus Subtest 2 (8,9%) sowie *Verwaltungsapparat* (6,9%) vorliegt. Es zeigt sich weiter, dass vor allem Probanden aus dem dritten und ersten Semester die Items nicht beantworten konnten (Anhang, Tabellen 13-15). Geistes- und Sozialwissenschaftler unterscheiden sich in ihrem Beantwortungsmuster nicht bedeutsam (Anhang, Tabellen 16-17). Orientiert an Lüdtke et. al. (2007) wurden fehlende Werte bis 5% als Fehler definiert. Für die Items *Lohnnebenkosten* und *Verwaltungsapparat*, für die die Analyse ein nicht-zufälliges, sondern scheinbar merkmalsabhängiges Fehlen zeigt, werden die fehlenden Werte über die Mittelwerte der jeweiligen Bezugsgruppe

(jeweiliges Semester und jeweiliger Studiengang) ersetzt, um die Stichprobengröße möglichst gut auszunutzen. Über ein Ausschlussverfahren würde neben einer Stichprobenverkleinerung der Mittelwert gegebenenfalls überschätzt werden. Imputationsverfahren wie bspw. die regressionsanalytische Schätzung der fehlenden Werte würden den Zusammenhang der Variablen, der anhand der Stichprobe erst exploriert werden soll, überschätzen.

10.5 Diskussion der Güte des Testverfahrens

Die Darstellung der testtheoretischen Überprüfung des Verfahrens zeigt, dass das der Hauptstudie zugrundeliegende Testverfahren im Hinblick auf die Schwierigkeit der Items eine große Bandbreite aufweist und nicht nur Items mittlerer Schwierigkeit umfasst. Damit kann einerseits sowohl eine Differenzierung zwischen Lösern und Nichtlösern erfolgen, andererseits auch eine Differenzierung unter Einbezug der jeweils itemspezifischen Schwierigkeiten. Um der selektiven Stichprobe der Studierenden gerecht werden zu können, wurden ebenfalls sehr schwierige Items konstruiert. Das Testverfahren will unterschiedliche Schwierigkeiten im Studienverlauf erfassen, die Komplexität des Konstrukts fächerübergreifender Problemlösekompetenzen explorieren und dennoch testökonomische Anforderungen erfüllen. Einbußen an Trennschärfe und Homogenität wurden deshalb hingenommen. Vor dem Hintergrund dieses Spannungsverhältnisses kann das Testverfahren im Hinblick auf die Gütekriterien der Objektivität, Reliabilität und Validität abschließend folgendermaßen bewertet werden:

Objektivität

„Ein Test ist dann objektiv, wenn er dasjenige Merkmal, das er misst, unabhängig von Testleiter, Testauswerter und von der Ergebnisinterpretation misst.“ (Moosbrugger & Kelava, 2007, S.9) Für das vorliegende Testverfahren kann festgestellt werden, dass durch die möglichst genaue Festlegung der Konzeption der Testunterlagen, die Testdarbietung und -instruktion, die Testauswertung sowie die Testinterpretation sowohl die Durchführungs- als auch Auswertungs- und Interpretationsobjektivität als gesichert gelten können. Um eine hohe Durchführungsobjektivität zu gewährleisten wurde ein standardisierter Testbogen ausgegeben, die Instruktion war ebenfalls standardisiert. Die Auswertungs- und Interpretationsobjektivität ist durch die Skalierung der Antwortvorgaben sowie die Bewertung der Aufgaben als richtig oder falsch eindeutig vorgegeben.

Reliabilität

Die Reliabilität von Cronbachs Alpha .31 als interne Konsistenz des Gesamttestverfahrens in der Hauptstudie ist vor dem besprochenen Hintergrund der Kürze und Dichotomie zu interpretieren

(Anhang, Tabelle 18). Reliabilitäten ab .7 sollten angestrebt werden, um zuverlässige Aussagen treffen zu können. Dabei spielt jedoch nicht nur die Homogenität und Testlänge eine bedeutsame Rolle, sondern auch die Varianz der Items, da Cronbach nicht nur davon ausgeht, dass alle Items dieselbe Zieldimension messen, sondern auch über die gleiche Varianz verfügen (Mossbrugger & Kelava, 2007). Die Reliabilität darf vor diesem Hintergrund für das vorliegende Testverfahren dementsprechend nur eine Komponente sein, die insbesondere durch die Beschränkung der Testlänge sowie die Heterogenität der Items geschwächt wird. Die Befunde sind vor diesem Hintergrund zu interpretieren.

Validität

Im Hinblick auf die Beurteilung der Validität kann auf unterschiedliche Validitätskonstrukte zurückgegriffen werden. So können unter anderem Augenschein- und Inhaltsvalidität sowie Kriteriums- und Konstruktvalidität zur Beurteilung des wichtigsten Gütemaßstabes herangezogen werden (Moosbrugger & Kelava, 2007). Im Rahmen der vorliegenden Studie konnte aufgrund des testökonomischen Vorgehens eine externe Validierung nicht dezidiert verfolgt werden, dennoch sollen im Folgenden einige Hinweise auf die Validität des Verfahrens dargestellt werden.

Mit Blick auf die Inhaltsvalidität zeigt die Darstellung der Konstruktionsprinzipien, dass sich das Testverfahren an der theoretischen Verortung des Konstrukts fächerübergreifender Problemlösekompetenz orientierte. Es wurden die dem Konstrukt zugrundeliegenden kognitiven Fähigkeiten sowie unterschiedliche Schwierigkeitsparameter bei der Konstruktion der Testaufgaben berücksichtigt. Im Hinblick auf die Beurteilung der Kriteriumsvalidität wird die Übereinstimmung des Konstrukts mit einem Außenkriterium herangezogen. Als mögliche externe Kriterien könnten für die vorliegende Studie sowohl Schulnoten als auch Intelligenztestwerte herangezogen werden. Dabei ist zu berücksichtigen, dass sich im vorliegenden Fall ein allgemeines Intelligenzmaß eher eignet als Schulnoten, die auf der Beurteilung domänenspezifischer Fähigkeiten, bzw. domänenspezifischem Wissen beruhen. Im Hinblick auf ein allgemeines Intelligenzmaß muss festgehalten werden, dass im Rahmen der vorliegenden Studie lediglich ein Intelligenzscreening durchgeführt wurde und die schwache Korrelation zwischen dem Gesamtscore der Testintelligenz und dem Gesamtscore fächerübergreifender Problemlösen lediglich als Indiz für einen möglichen Zusammenhang gelten kann ($r = .16, p < .01$). Dezidiert wird auf Zusammenhänge zwischen Intelligenz und Komponenten fächerübergreifender Problemlösekompetenz bei der Darstellung der Befunde eingegangen.

Insgesamt bleibt festzuhalten, dass die Befunde vor dem Hintergrund der dargestellten Güte des Testverfahrens sowie mit Blick auf die ausgewählte Stichprobe zu interpretieren sind. Im Rahmen der abschließenden Diskussion der Studie wird auf Möglichkeiten der Erweiterung und Optimierung des Testinstrumentes hingewiesen.

11. Beschreibung der Hauptstudie

Im Folgenden werden Untersuchungsdesign, Stichprobe sowie Datenerhebungszeiträume der Hauptstudie vorgestellt.

11.1 Übersicht über Testaufgaben und Items

Die nachstehende Tabelle fasst nochmals die der Hauptstudie zugrundeliegenden Subtests sowie deren Itemsets und Einzelitems zusammen.

Grafik 7: Items die in die Hauptstudie aufgenommen wurden

Subtest 1	Situationsanalyse	Induktives Schlussfolgern	Handwerker
			Migrationsbiographie
		Deduktives Schlussfolgern	Berufstätige Frauen
			Familienunternehmen
			Gesundheitskosten
		Kausalität und Korrelation	Verkehrsunfälle
			Alkoholkonsum
		Fakten und Meinung	Arbeitslose
Jugendliche Raucher			
Subtest 2	Zielevaluation	Bewertung und Einschätzung von Zielen	Verwaltungsapparat
			Lohnnebenkosten
Subtest 3	Handlungsplanung	Reihung und Platzierung von Handlungsschritten	Arbeitslosigkeit
			Getränkehersteller

10.2 Untersuchungsdesign

Die Studie wurde als quasi-experimentelle Feldstudie im Querschnittsdesign auf Ebene des Bachelorstudiums angelegt. Ausgewählt wurde das Untersuchungsdesign aufgrund der Implikationen der Fragestellung: Die Frage nach der fächerübergreifenden Problemlösekompetenz Studierender erfordert es, die zu analysierenden Kompetenzen nicht nur über ein Testszenario abzufragen, das berufliches fächerübergreifendes Problemlösen simuliert. Die Studie sollte ebenfalls auf differentialpsychologischer Ebene Einblick in die Kompetenzausprägungen bei Studierenden erlauben. Die Studie fokussiert dementsprechend darauf, Studierende als „natürliche Gruppen“ zu untersuchen und zu vergleichen. Der damit einhergehende Nachteil der geringeren internen Validität wurde über die Kontrolle möglicher konfundierender Variablen versucht auszugleichen. So wurden einerseits bedeutsame personbezogene Kovariaten kontrolliert und weiterhin einer sozial erwünschten Antworttendenz über die Angabe von Codewörtern sowie über die Zusicherung der Anonymität

entgegengewirkt. Andererseits wurden externe Störvariablen über eine ausführliche und klare Instruktion sowie eine Beschreibung des Untersuchungsverlaufs kontrolliert. Der Rahmen eines Querschnittsdesigns wurde gewählt, um unterschiedliche Studierendengruppen kontrastieren zu können und Kompetenzausprägungen im Querschnitt zu untersuchen. Matchingverfahren oder Längsschnittanalysen sind schwer zu realisieren. Die Eingeschränktheit von Querschnittanalysen in Bezug auf die Untersuchung von Entwicklungsverläufen und Kausalzusammenhängen muss hier berücksichtigt werden. Die in den Befunden dargestellte Kontrastierung fächerübergreifender Problemlösekompetenzen bei unterschiedlichen Studierendengruppen sowie die Beschreibung von Zusammenhängen ist vor dem Hintergrund dieser Einschränkung zu sehen.

11.3 Stichprobe

Die Zielstichprobe waren Studierende verschiedener Studienfächer und unterschiedlicher Semester auf Bachelorebene. Aufgrund der Zugänglichkeit wurden Studierende des ersten, dritten und sechsten Semesters im Bachelorstudium an der Katholischen Universität Eichstätt-Ingolstadt ausgewählt. Eine Anfrage an weiteren Universitäten zur Erweiterung des Fächerspektrums auf den naturwissenschaftlichen Bereich war nicht erfolgreich. Aufgrund der kontrastierend ausgerichteten Anlage der Studie wurden zwei möglichst dichotome Studiengruppen, Sozial- und Geisteswissenschaftsstudierende, ausgewählt. Befragt wurden alle Erst-, Dritt- und Sechstsemesterstudierenden der Studiengänge Lehramt Schulpsychologie und Bachelor Journalistik – für die Sozialwissenschaften, sowie der Studiengänge des Lehramtes- bzw. Bachelors Anglistik und Bachelor Europastudien – als geisteswissenschaftliche Kohorte. Die Studierenden wurden über ihre Dozenten rekrutiert, ihre Teilnahme war freiwillig. Die Untersuchung wurde im Rahmen der jeweiligen Veranstaltung durchgeführt, um eine möglichst hohe Beteiligungsquote zu erreichen. Die Testung fand am Ende der Veranstaltung statt, sodass gewährleistet wurde, dass Studierende, die nicht an der Studie teilnehmen wollten, die Veranstaltung verlassen konnten. Zwei Studierende nahmen diese Möglichkeit wahr.

11.4 Untersuchungsverlauf

Die Testzeit für die Durchführung des Testverfahrens sowie eine Bearbeitung der begleitenden Fragebögen und des Intelligenzscreenings umfasste zirka 1,5 Stunden. Dementsprechend wurde die Testzeit auf zwei Tage aufgeteilt: Am ersten Testtag wurde der Problemlösetest durchgeführt, am zweiten Testtag wurden die Fragebögen sowie das Intelligenzscreening ausgegeben. Das Testverfahren zum fächerübergreifenden Problemlösen erfordert ca. 45 Minuten Testzeit – die genaue Testzeit der Probanden wurde nicht erhoben, es war jedoch gewährleistet, dass jeder Testteilnehmer ausreichend Zeit zur Bearbeitung zur Verfügung hatte. Je nach zeitlichen Möglichkeiten der

Lehrenden wurden je ein Testteil an zwei Testtagen zu je zirka 45 Minuten durchgeführt oder – in zwei Fällen – beide Testteile im Rahmen einer mehrstündigen Veranstaltung durchgeführt, wobei auch hier die genannte Reihenfolge beibehalten blieb. Der Testzeitpunkt fand in Absprache mit den Lehrenden statt, weshalb die Tageszeit, zu der die Testung stattfand, nicht selbst bestimmt werden konnte. Insgesamt wurde die Hauptstudie an 22 Testtagen durchgeführt, im Rahmen derer insgesamt 13 Studierendengruppen entweder eineinhalbstündig oder zweimal fünfundvierzigminütig getestet wurden. Die Studierenden wurden für ihre Unterstützung mit Süßigkeiten belohnt. Zudem hatten sie die Möglichkeit, sich auf eine von den Testheften getrennte E-Mailliste einzutragen, um nach der Auswertung der Gesamtergebnisse über die aggregierten Studienergebnisse informiert zu werden. Einzelergebnisse wurden nicht herausgegeben und ausgewertet.

Der Untersuchungsverlauf war für alle Testzeitpunkte standardisiert: So wurden bereits vor Beginn der Vorstudie Lehrende aus den ausgewählten Studiengängen und Semestern angefragt, ob sie mit ihren Studierenden an der Studie teilnehmen möchten. Die Dozenten informierten ihre Studierenden über die Durchführung der Studie, es wurde ein geeigneter Testzeitpunkt gefunden. Es wurde vor Beginn der Testung eine mündliche Testanweisung gegeben. Die Intention der Studie wurde vorgestellt, es wurde erläutert, welche Inhalte die Studierenden erwarten. Vor Ausgabe der Testhefte wurde darauf hingewiesen, dass sich die Studierenden in die im Testheft beschriebene Situation versetzen sollten. Die Studierenden wurden gebeten, die Aufgabenstellungen möglichst zügig und vollständig zu bearbeiten und vor Abgabe des Testheftes die Vollständigkeit der Aufgabenbearbeitung nochmals zu überprüfen. Es wurde erläutert, dass alle Aufgaben lösbar sind und die Bearbeitung der Testaufgaben nochmals explizit im Testheft erläutert wird. Nach der Bearbeitung wurden die Testhefte eingesammelt, sobald alle Studierenden ihre Bearbeitung beendet hatten. Da die Studierenden jederzeit die Möglichkeit hatten, den Test abbrechen, dies aber niemand in Anspruch nahm, kann, obwohl nicht explizit erhoben, davon ausgegangen werden, dass die Motivation sich am Testverfahren zu beteiligen ausreichend gegeben war.

11.5 Untersuchungszeitraum

Die Teilnehmer an der Hauptstudie wurden im Sommersemester 2010 rekrutiert. Die Pilotierung fand zwischen November und Dezember 2010 statt. Die Hauptstudie fand von Januar bis Juni 2011 statt.

Grafik 8: Untersuchungszeitraum



11.6 Varianz der Hauptstichprobe

Die nachfolgend dargestellten Daten zur Größe der Hauptstichprobe und zur Verteilung von Alter, Geschlecht, Bildungserfahrung oder auch Bildungshintergrund stellen die Verteilungsmuster innerhalb der Stichprobe vor. Repräsentative Aussagen können nicht erfolgen.

Studierende, Studiengänge und Semesterverhältnis

Im Rahmen der Studie wurden insgesamt N=290 Studierende getestet. Nach Studiengängen aufgegliedert zeigt sich, dass sowohl aus dem Bereich der Geisteswissenschaften als auch dem Bereich der Sozialwissenschaften annähernd gleich viele Studierende teilnahmen. Eine ausgeglichene Fallzahl wurde zwar von vornherein angestrebt, konnte jedoch aufgrund nicht kontrollierbarer Einflussfaktoren (unterschiedliche Zulassungszahlen in den Studiengängen, Abbrecherquoten, unterschiedliche Veranstaltungsgrößen) nicht festgelegt werden.

Tab. 1: Teilnehmer nach Studiengängen

		Semester			Gesamt
		1,00	3,00	6,00	
Studiengang	Bachelor of Arts	24	13	17	54
	Europastudien				
	Anglistik im Rahmen des Lehramtes	31	20	29	80
	Schulpsychologie im Rahmen des Lehramtes	22	20	32	74
	Bachelor of Arts	39	19	24	82
	Journalistik				
Gesamt		116	72	102	290

Mit N=156 Studierenden aus dem Bereich der Sozialwissenschaften (Journalistik: N=82 und Schulpsychologie N=74) und N=134 Studierenden aus dem Bereich der Geisteswissenschaften (Europastudien: N=54 und Anglistik: N=80) ist sowohl die Gesamtzahl der Teilnehmer als auch die Ausgewogenheit gut.

Der Blick auf die Semester zeigt, dass das dritte Semester mit N=72 Studierenden gegenüber N=116 Erstsemestern und N=102 Sechstsemestern unterrepräsentiert ist. Eine Aufgliederung nach Studiengängen zeigt, dass zwar wiederum eine annähernde Gleichverteilung zwischen den Studiengängen erreicht werden konnte, allerdings der Europastudiengang im dritten Semester

unterrepräsentiert ist. Dies lag daran, dass viele Studierende des dritten Semesters des Europastudiengangs im Rahmen einer längeren Exkursion im Ausland waren.

Geschlechterverhältnis

Nach Geschlecht aufgeteilt zeigt sich, dass sich auch im Rahmen der Studie das Geschlechterverhältnis einer frauendominierten Universität widerspiegelt: Es nahmen N=204 Frauen und N=57 Männer an der Studie teil. N=29 Studierende machten keine Angaben.

Tab. 2: Teilnehmer nach Geschlecht und Studiengang

		Geschlecht		Gesamt
		weiblich	männlich	
Studiengang	Bachelor of Arts Europastudien	48	3	51
	Anglistik im Rahmen des Lehramtes	53	16	69
	Schulpsychologie im Rahmen des Lehramtes	56	13	69
	Bachelor of Arts Journalistik	47	25	72
Gesamt		204	57	261

Nach Studiengängen aufgegliedert zeigt sich, dass die Teilnahme der weiblichen Teilnehmer relativ ausgeglichen ist. Im Europastudiengang nahmen die wenigsten, im Bereich der Journalistik die meisten Männer an der Studie teil.

Studierende: Alter, Bildungserfahrung und Bildungshintergrund

Das Durchschnittsalter der Teilnehmer der Hauptstudie lag bei 22,01 Jahren (SD = 2,89), der oder die jüngste Teilnehmer, bzw. Teilnehmerin war 18, der oder die älteste Teilnehmer, bzw. Teilnehmerin 46 Jahre. Ungefähr drei Viertel der Teilnehmer waren zwischen 20 und 24 Jahren alt. Die durchschnittliche Abiturnote lag bei $\bar{x}=2,10$ (SD= .55). N=19 Studierende gaben an, vor diesem Studium bereits ein anderes Studienfach belegt zu haben (N=3 Europastudien, N=8 Anglistik, N=4 Schulpsychologie, N=4 Journalistik). N=27 Studierende gaben diesbezüglich keine Auskunft.

Eine Berufsausbildung hatten bereits N=20 Studierende absolviert (N=4 Europastudien, N=9 Anglistik, N=5 Schulpsychologie, N=2 Journalistik). N=33 Studierende gaben diesbezüglich keine Auskunft. Praktikumserfahrung sammelten fast alle Studierenden (N=244), die Studierenden der Journalistik konnten die meiste Praktikumserfahrung nachweisen (N=13 Studierende mit ≥ 2 Jahren Praktikumserfahrung). Drei Viertel der Studierenden (76%) schätzen sich auf einer Skala von 1 („keine berufliche Erfahrung“) bis 4 („beruflich sehr erfahren“) zum Testzeitpunkt insgesamt als

„wenig“ bis „nicht beruflich erfahren“ ein (N=27 fehlend). Die subjektiv empfundene Berufserfahrung nimmt bei einem Vergleich der Semester signifikant zu (1. Semester $\bar{x}=1,87$; 3. Semester $\bar{x}=2,13$; 6. Semester $\bar{x}=2,23$; $F=7,29$, $df=152,71$, $p<.01$). Die Daten zum Bildungshintergrund der Eltern zeigen, dass der überwiegende Teil der Studierenden (N=199) aus einem Elternhaus kommt, in dem mindestens ein Elternteil Abitur oder einen höheren Bildungsabschluss (Studium oder Promotion) erreicht hat (Anhang, Tabelle 25).

11.7 Beschreibung der weiteren Verfahren

Im Folgenden werden diejenigen Verfahren beschrieben, die im Rahmen der Hauptstudie begleitend eingesetzt wurden. Die Bearbeitungszeit der Verfahren nahm zwischen 30 bis 45 Minuten in Anspruch. Die Fragebögen und das Intelligenzscreening wurden von den Studierenden am zweiten Testtag bearbeitet, bzw. als zweiter Testteil, sofern beide Testteile an einem Tag durchgeführt wurden. Die Skalen zu Studieninteresse, epistemologischen Überzeugungen und Selbstwirksamkeit wurden im Rahmen eines Testheftes zusammengefasst. Begleitend wurden die Matrizen des Intelligenzscreenings sowie der entsprechende Antwortbogen ausgegeben. Dem Testheft mit den Fragebogenskalen sowie dem Matrizenbogen lagen eine Instruktion und eine Aufgabenbeschreibung bei. Die Studierenden wurden eingangs mündlich instruiert. Die begleitend eingesetzten Fragebögen sind im Anhang angefügt.

Studieninteresse

Die Skala zum Studieninteresse basiert auf dem Fragebogen zum Studieninteresse (FSI) von Krapp, Schiefele, Wild und Winteler (1993; Anhang Tab. 19). Interesse an Lerninhalten ist eine wichtige Bedingung erfolgreicher Lernprozesse (vgl. Kap. 5.5). Interesse als motivationaler Faktor ist dementsprechend bedeutungsvoll für das Studium. Die Autoren definieren Interesse formal als spezifische Relation zwischen einer Person und einem Gegenstand, im vorliegenden Fall dem Studienfach. Krapp, Schiefele, Wild und Winteler (1993) konnten bei der Überprüfung ihres Fragebogens zeigen, dass Studieninteresse mit intrinsischer motivationaler Orientierung, Tätigkeitszentrierung und studienbezogener Autonomieorientierung korreliert. Sie stellen weiter dar, dass die Interessenkomponenten *gefühlsbezogene* und *wertbezogene Valenzen* und *intrinsischer Charakter* nicht als unabhängige Faktoren aufgefasst werden können. Für die vorliegende Studie wurden aus dem ursprünglich 27 Items umfassenden FSI je zwei Items aus den Dimensionen „Persönliche wertbezogene Valenzen“, „Intrinsischer Charakter“ und „Gefühlsbezogene Valenzen“ entnommen. Die Skala umfasst dementsprechend sechs Items (Anhang, Tabelle 19). Die interne Konsistenz der ursprünglichen Skala liegt bei Cronbachs Alpha .90. Die hier verwendete gekürzte Skala weist eine interne Konsistenz von Cronbachs Alpha .77 auf. Es lag ein vierstufiges

Antwortformat von „Trifft nicht zu“ (1), „Trifft kaum zu“ (2), „Trifft eher zu“ (3), „Trifft genau zu“ (4) vor.

Selbstwirksamkeit

Die Skala zur studienbezogenen Selbstwirksamkeit basiert auf der Skala der schulbezogenen Selbstwirksamkeit von Jerusalem & Satow (1999; Anhang Tab. 20). Sie umfasst insgesamt 7 Items. Der in der ursprünglichen Version verwendete Schulbezug wurde in einen Studienbezug umformuliert. Die studienbezogene Selbstwirksamkeitserwartung fokussiert die Kompetenzerwartungen von Studierenden im Umgang mit den jeweiligen Anforderungen ihres Studiums (vgl. Kap. 5.5). Die interne Konsistenz der ursprünglichen Skala liegt zwischen Cronbachs Alpha .70 und .73. In der vorliegenden Studie liegt die interne Konsistenz bei Cronbachs Alpha .71. Es lag ein vierstufiges Antwortformat von „Trifft nicht zu“ (1), „Trifft kaum zu“ (2), „Trifft eher zu“ (3), „Trifft genau zu“ (4) vor.

Epistemologische Überzeugungen

Vorstellungen und Überzeugungen der Studierenden im Zusammenhang mit dem jeweiligen fachwissenschaftlichen Studium beeinflussen die Vorstrukturierung der wahrgenommenen Inhalte des Studiums und nehmen Einfluss auf das Lernen der Studierenden (vgl. Kap. 5.5). Im Zentrum des Interesses standen für die Studie folgende Überzeugungen:

- Wie schätzen Studierende die *Relevanz fächerübergreifender Probleme im späteren beruflichen Kontext* ein? Gehen Studierende davon aus, dass sie im späteren Beruf ausschließlich mit domänenspezifischen Problemen konfrontiert werden? Oder sind sie davon überzeugt, dass sie mit Problemen konfrontiert werden, die auch weniger mit ihrem Studium zu tun haben können? (Anhang, Tabelle 21)
- Wie schätzen Studierende den *Transfer fachspezifisch erlernter Problemlösemethoden auf fachübergreifende Problemstellungen* ein? Sind Studierende der Überzeugung, dass sich ihre fachspezifisch erworbenen Problemlösekompetenzen auf fächerübergreifende Probleme übertragen lassen? Oder gehen sie eher davon aus, dass sie mittels ihrer fachspezifischen Methoden, Strategien und Problemlöseansätzen lediglich fachspezifische Probleme lösen können? (Anhang, Tabelle 22)
- Wie schätzen Studierende die *Quelle ihres jeweiligen fachspezifischen Wissens* ein? Gehen Studierende davon aus, dass sie den Autoritäten fachwissenschaftlichen Wissens immer Glauben schenken müssen? Oder hinterfragen Studierende die Inhalte fachwissenschaftlicher Literatur, bzw. ihrer Dozenten? (Anhang, Tabelle 23)

- Wie schätzen Studierende die *Sicherheit ihres fachspezifischen Wissens* ein? Gehen Studierende davon aus, dass sich das fachliche Wissen immer weiter entwickeln wird oder sind sie eher der Annahme, dass das Wissen in ihrem Fach festgelegt ist und nichts Neues hinzukommen oder bereits Etabliertes ggf. revidiert wird? (Anhang, Tabelle 24)

Zur Einschätzung der Relevanz fächerübergreifender Probleme im späteren beruflichen Kontext wurden insgesamt 4 Items konzipiert. Ziel war es, abzufragen, wie hoch die Relevanz fachübergreifender Probleme im späteren Beruf eingeschätzt wird. Erst wenn fachübergreifende Probleme für den späteren Beruf überhaupt als Aufgabe angesehen werden, kann davon ausgegangen werden, dass eine entsprechende Motivation zur Beschäftigung mit diesen im Studium vorliegt. Die Skala *Relevanz* weist eine interne Konsistenz von Cronbachs Alpha .54 auf.

Zur Einschätzung der Transfermöglichkeiten wurden insgesamt 5 Items entwickelt. Die Items sollten abfragen, inwieweit Studierende davon ausgehen, dass die von ihnen im Rahmen des Fachstudiums erlernten Problemlösemethoden ihnen auch bei der Lösung fachfremder Probleme helfen. Abgefragt wurde die Nützlichkeit fachspezifischer Herangehensweisen bei der Analyse von fachfremden Problemen, um ein Problemverständnis zu entwickeln, die Nützlichkeit fachspezifischer Herangehensweisen zur Entwicklung von Lösungsmöglichkeiten bei fachfremden Problemen, die Nützlichkeit fachspezifischer Methoden zur strategischen Lösung von fachfremden Problemen und die Nützlichkeit fachspezifischen Wissens zur Einordnung von fachfremden Problemen. Die Skala *Transfer* weist eine interne Konsistenz von Cronbachs Alpha .80 auf.

In Anlehnung an den von Urhahne und Hopf (2004) entwickelten Fragebogen zur Einschätzung epistemologischer Überzeugungen im naturwissenschaftlichen Bereich wurden die epistemologischen Überzeugungen der Studierenden im Hinblick auf das Wissen in ihrem Studienfach abgefragt. Die Autoren konnten zeigen, dass reflektiertere epistemologische Überzeugungen mit einer höheren Leistungsmotivation und einem höheren fachspezifischen Selbstkonzept einhergehen. Die Überzeugung eher statischen Wissens geht hingegen mit der Bevorzugung von einfachen Memorierstrategien einher. Die Überzeugung von der Veränderlichkeit des Wissens und die Überzeugung von Autoritäten als Wissensvermittlern geht einher mit kognitiv aufwändigeren Lernstrategien (Urhahne & Hopf, 2004). Für die vorliegende Studie wurden die beiden Dimensionen *Sicherheit des Wissen* und *Quelle des Wissens* verwendet. Die von den Autoren auf den naturwissenschaftlichen Bereich ausgerichteten Items wurden in ungekürzter Anzahl übernommen, jedoch auf das jeweilige Studienfach bezogen umformuliert (Sicherheit des Wissens: „Das Wissen in der Fachwissenschaft, die ich studiere, ist für alle Zeit wahr“; Quelle des Wissens: „Was in Büchern meiner Fachwissenschaft steht, muss man glauben“). Die Skala *Sicherheit des Wissens* umfasst insgesamt 6 Items. In der Originalversion weist sie eine interne Konsistenz von Cronbachs Alpha .41 auf, in der umformulierten Version eine interne Konsistenz von Cronbachs Alpha .68. Die Skala *Quelle des Wissens* umfasst insgesamt 5 Items. In der von Urhahne und Hopf konzipierten Skala weist

sie eine interne Konsistenz von Cronbachs Alpha .67, in der hier vorliegenden umformulierten Version eine interne Konsistenz von Cronbachs Alpha .80. Allen Skalen lag ein vierstufiges Antwortformat von „Trifft nicht zu“ (1), „Trifft kaum zu“ (2), „Trifft eher zu“ (3), „Trifft genau zu“ (4) zugrunde.

Begleitend wurden demographische Variablen abgefragt (Geschlecht, Alter, Abiturnote, Erststudium, Bildungshintergrund der Eltern, Berufsausbildung, Praktikumserfahrung in Monaten, Einschätzung der beruflichen Erfahrung; vgl. Fragebogen im Anhang).

Intelligenzscreening

Zur Kontrolle der Intelligenz wurde das Set I des *Advanced Progressive Matrices Tests* (APM) von Raven (dt. Bearbeitung von Bulheller & Häcker, 1998) verwendet (vgl. Kap. 5.5). Die sprach- und kulturfreien Testaufgaben des APM bestehen aus unvollständigen geometrischen Figuren oder Mustern, sogenannten Matrizen, die aus jeweils acht Antwortalternativen ergänzt werden sollen. Die Aufgaben beziehen sich auf analytisches, schlussfolgerndes und kombinatorisches Denken, das auf figürliches Material angewendet werden muss. Das Set umfasst insgesamt 12 Matrizen in aufsteigender Schwierigkeit und kann bei Jugendlichen und Erwachsenen ab 15 Jahren sowohl als Einzel- als auch Gruppentest durchgeführt werden. Retestreliabilitäten liegen zwischen .76 und .91, die interne Konsistenz liegt zwischen .83 und .87. Neben Korrelationen zu anderen Intelligenztests (bspw. HAWIE, zwischen $r=.25$ und $r=.50$) liegen Werte zur Vorhersagevalidität im beruflichen Setting vor. So wurde bei einer Gruppe von 1.120 Managern ein Zusammenhang von $r=.20$ zwischen der Testleistung im APM und der Leistungsbeurteilung im Unternehmen nachgewiesen (Bulheller & Häcker, 2005). Da Set I und Set II zusammen etwa 1 Stunde Bearbeitungszeit in Anspruch nehmen, Set II allein etwa 40 Minuten, wurde Set I als Screeninginstrument im Rahmen der Hauptstudie eingesetzt.

Teil IV: Befunde der Hauptstudie

Im Folgenden werden die Befunde der Hauptstudie dargestellt, orientiert an den Fragenkomplexen und Hypothesen: Zunächst wird danach gefragt, inwieweit sich das im Theorieteil beschriebene Konstrukt fächerübergreifenden Problemlösens über das Testverfahren empirisch nachzeichnen lässt. Anschließend wird erläutert, inwieweit es Zusammenhänge zwischen den erhobenen Kovariaten und fächerübergreifender Problemlösefähigkeit gibt. Im darauffolgenden Abschnitt werden die Kompetenzen fächerübergreifenden Problemlösens bei Studierenden beschrieben. Abschließend werden die Befunde zusammengefasst.

12. Verteilung der Daten

Es wurde geprüft, inwieweit die Daten den Anforderungen der Normalverteilung und Varianzhomogenität entsprechen. Zur Beurteilung der Normalverteilung wurden verschiedene Anhaltspunkte herangezogen. Schiefe und Exzess sowie Signifikanztests für große Stichproben (Kolmogorov-Smirnov) und kleinere Gruppen (Shapiro-Wilk) wurden sowohl für die Gesamtstichprobe als auch für unterschiedliche Semester und Disziplinen verglichen (Anhang, Tabelle 26- 27).

Bereits bei der Konzeption der Studie wurde über eine möglichst große Stichprobe versucht, Standardfehler gering zu halten. Größere Stichproben führen jedoch auch dazu, dass bereits geringe Abweichungen von der Normalverteilung signifikant werden (Bortz, 2005). Zur deskriptiven Beurteilung der Daten wurden deshalb auch QQ-Diagramme und Histogramme herangezogen (nicht im Anhang enthalten). Die Daten zeigen, dass zwar nicht von einer Normalverteilung ausgegangen werden kann, auffällige Abweichungen wie extrem schiefe oder mehrgipflige Verteilungen allerdings nicht vorliegen (Anhang, Tab. 26). Von einer einheitlichen Grundgesamtheit im Hinblick auf Disziplinen und Semester kann ausgegangen werden, eine Ausnahme bildet das Itemset „Fakten und Meinung“ (Kolmogorov-Smirnov-Z; Anhang, Tab. 27).

Mit Blick auf Varianzanalysen steht die Frage der Varianzhomogenität im Vordergrund (für einen Überblick Diehl & Arbinger, 1990). Die Daten zeigen, dass von homogenen Varianzen ausgegangen werden kann (Levene-Tests), eine Ausnahme bildet das Itemset „Fakten und Meinung“. Vor dem Hintergrund der Verteilung der Daten wird bei der Berechnung, sofern notwendig, auch auf nicht-parametrische Unterschiedsanalysen zurückgegriffen (Mann-Whitney-U, Kruskal-Wallis). Für weitere Berechnungen wurden die Daten z-transformiert, um unterschiedlich verteilte und bepunktete Daten miteinander vergleichen zu können.

13. Fragenkomplex I: Konstruktklärung

Der Fragenkomplex prüft, inwieweit fächerübergreifendes Problemlösen als mehrdimensionales Konstrukt empirisch nachgezeichnet werden kann. Die Heuristik der Problemlösephasen wurde für das Testverfahren über die Subtests *Situationsanalyse*, *Zielevaluation* und *Handlungsplanung* operationalisiert. Es wurden Fähigkeiten des logischen Schließens, Evaluierens, Vergleichens und Planens abgebildet. Insgesamt wurden 13 Items unterschiedlicher Schwierigkeitsstufen konstruiert. Subtest 1 umfasst neun Items, die Subtests 2 und 3 jeweils zwei Items. Im Folgenden wird danach gefragt, inwieweit die manifesten Variablen gemeinsame Varianz teilen und inwieweit (Sub)Dimensionen und Binnenstrukturen des Konstruktes identifiziert werden können.

Für den Fragenkomplex wurden folgende Hypothesen formuliert.

- Über die Subtests lassen sich kognitive Teilkompetenzen differenzieren, die der Heuristik des Problemlöseprozesses zugeordnet werden können (Hypothese 1).
- Über die Itemsets lassen sich voneinander unterscheidbare Teilkompetenzen des deduktiven, induktiven und korrelativen Schließens sowie der Unterscheidung von Fakten und Meinung unterscheiden (Hypothese 2).
- Fächerübergreifende Problemlösekompetenzen und Testintelligenz teilen in Abhängigkeit der zugrundeliegenden Anforderungsstruktur des Problems gemeinsame Varianz (Hypothese 3).
- Studierende der Sozialwissenschaften und Studierende der Geisteswissenschaften unterscheiden sich im Hinblick auf ihre fächerübergreifende Problemlösekompetenz. Es wird angenommen, dass Studierende der Sozialwissenschaften bessere Leistungen bei der Lösung von Problemen zeigen, die auf induktivem und kausalem Schließen basieren (Hypothese 4).

13.1 Verwendete Verfahren

Zur Überprüfung der Hypothesen wurde exploratorisch die Faktorstruktur des Konstruktes geprüft. Über konfirmatorische Faktorenanalysen wurde außerdem die Validität der Faktormodelle getestet. Für die Faktorenanalyse wurden folgende Festlegungen getroffen (für einen Überblick u.a. Moosbrugger & Kelava, 2008; Diehl & Staufenbiel, 2007; Backhaus, Erichson, Plinke & Weiber, 2006; Bortz & Döring, 2006): Die exploratorische Faktorenanalyse wurde als Hauptkomponentenanalyse durchgeführt, um die gesamte Varianz der Variablen über eine ordnende Faktorstruktur zu erklären und nach einem „Sammelbegriff“ für die Variablenbündel zu suchen. Es wird für eine erste Analyse gemeinsamer Varianz angenommen, dass der gesuchte Faktor über die Variablen restlos erklärt werden kann. Begleitend wurde über eine Hauptachsenanalyse die Spezifität der Variablenvarianz sowie die Messfehlervarianz berücksichtigt. Als Kriterien wurden folgende Richtmaße zugrunde gelegt: Das Kaiser-Meyer-Olkin-Kriterium, das einen Wert $> .5$ aufweisen sollte, der Bartlett-Test auf Sphärizität, der auf die Annahme der Alternativhypothese und die

Normalverteilung der Daten hinweisen sollte, der Scree-Test, der die Faktoren und ihre Eigenwerte in abfallender Reihenfolge abträgt, die Höhe der Faktorladungen und ebenfalls die Kommunalitäten der einzelnen Variablen sowie die über die Faktoren aufgeklärte Varianz. Das Eigenwertkriterium wurde als ausschlaggebendes Kriterium herangezogen, wobei die Eigenwerte > 1 sein sollten, das heißt, der extrahierte Faktor sollte mehr Varianz klären als eine einzige Variable. Angeführt wurden ebenfalls Faktoren, die zwar das Eigenwertkriterium verletzen, jedoch noch mehr als 10% der Gesamtvarianz klären. Die Faktorladungen wurden über Rotationsmethoden transformiert, wobei sowohl unkorrelierte (orthogonale Varimax-Rotation) als auch korrelierte (direktes Oblimin) Transformationen zur Interpretation herangezogen wurden.

Die konfirmatorische Faktorenanalyse (CFA) auf Basis einer Maximum-Likelihood-Schätzung (ML) wurde zur Struktur- und Modellprüfung herangezogen, um die exploratorisch indizierten Zusammenhänge zwischen den hypothetischen Konstrukten und den beobachteten Indikatorvariablen nochmals zu überprüfen (für einen Überblick Backhaus et. al., 2006; Bühner, 2006; Beauducel & Wittmann, 2005; Hu & Bentler, 1999). Die Stichprobe erfüllt die grundlegenden Voraussetzungen einer hinreichend großen Stichprobe von mindestens $N > 100$, bzw. besser $N > 200$ und mindestens $N > 5 \cdot t$, wobei t die Anzahl der zu schätzenden Parameter darstellt. Dabei setzt die CFA für eine Schätzung nach dem Maximum-Likelihood-Prinzip normalverteilte Daten voraus, die Schätzer sind jedoch relativ robust gegenüber Abweichungen. Um den Modellfit zu bestimmen, können folgende Kriterien als Kennwerte für eine gute Modellpassung gelten: Als gebräuchlichster Modellfitwert im Rahmen der ML-Schätzung wird der *Root Mean Square Error of Approximation* (RMSEA) herangezogen, der die Abweichung zwischen den durch das Modell postulierten und den tatsächlich erhobenen Daten ermittelt. Diese Diskrepanz wird anschließend durch die Anzahl der Freiheitsgrade dividiert. Auf einen guten Modellfit weisen Werte zwischen .00 und .05 hin, akzeptable Werte liegen zwischen .051 und .085. Ist RMSEA größer als .1 spricht dies gegen das Modell. Daneben kann auch das Maß des *Comparative Fit Index* (CFI) herangezogen werden. Da der CFI bei Verletzungen der Normalverteilung dazu neigt gering auszufallen (Beauducel & Wittmann, 2005), wird begleitend der *Normed Fit Index* (NFI) verglichen, der jedoch sehr sensibel auf die Stichprobengröße reagiert. Beide Werte sollten im optimalen Fall Werte von 1.0 aufweisen, für einen guten Modellfit sollten sowohl der CFI als auch der $NFI \geq .9$ sein (Mossbrugger & Kelava, 2008; Backhaus et. al., 2006). Der Tucker-Lewis Index (TLI), als Maß der Übereinstimmung zwischen einem hypothetischen und gemessenen Modell sollte ebenfalls bei $> .9$ liegen. Für die Chi-Quadrat-Werte sollte $\chi^2 < 2 \text{ df}$ gelten (Backhaus et. al., 2006). Über AMOS kann das Maß über ‚CMIN/DF‘ abgelesen werden, wobei das Modell als umso besser gilt, je kleiner χ^2/df ist. Es liegt jedoch kein Konsens vor, wann ein ‚guter‘ Modellfit erreicht ist, die Empfehlungen schwanken zwischen 3,2 oder kleiner bis hin zu 5 (Arbuckle & Wothke 1999; Bollen, 1989). Der Modelltest sollte nicht signifikant ausfallen, wobei eine größere Stichprobe sowie eine Verletzung der Normalverteilung leicht zu einer Signifikanz führen können. In diesem Fall

sollten die Fit-Indizes stärker gewichtet werden (Bühner, 2006). Insgesamt wird empfohlen, den Modellfit über mehrere Kennwerte zu berichten, Hair et. al. (2006) empfehlen drei Indices.

13.2 Binnenstruktur des Konstruktes

Ausgehend von der Annahme, dass die drei Subtests *Situationsanalyse*, *Zielevaluation* und *Handlungsplanung* Teilkomponenten fächerübergreifender Problemlösekompetenz sind, wird geprüft, inwieweit die manifesten Variablen auf einem gemeinsamen latenten Faktor laden. Die exploratorische Faktorenanalyse zeigt, dass über ein einfaktorielles Modell 43% der Gesamtstreuung geklärt werden können (Anhang, Tabelle 28-29). Die Subtests laden mit .40 (*Situationsanalyse*) und .73 (*Zielevaluation*) sowie .77 (*Handlungsplanung*) auf dem gemeinsamen Faktor (Anhang, Tabelle 28). Der gemeinsame Faktor erklärt allerdings für die Subtests *Zielevaluation* und *Handlungsplanung* mehr Varianz als für den Subtests *Situationsanalyse* (Anhang, Tabelle 28-30). Der Vergleich der korrelierten und unkorrelierten Faktorenlösungen zeigt, dass die Subtests *Zielevaluation* und *Handlungsplanung* über einen gemeinsamen Faktor *bewertendes und planendes Schließen* von kognitiven Fähigkeiten der *Situationsanalyse* getrennt betrachtet werden müssen. Allerdings kann nicht von einer Null-Korrelation zwischen den Faktoren ausgegangen werden.

Können unterschiedliche Teilkomponenten logischen Schließens empirisch differenziert werden? Eine Analyse der Itemsets *deduktives Schließen*, *induktives Schließen*, *korrelatives Schließen*, *subjektive und objektive Meinung unterscheiden*, *Handlungen planen*, *Ziele evaluieren* bestätigt die Faktorstruktur auf Subtestebene. Es zeigt sich, dass zwei Faktoren extrahiert werden können, die mit Kommunalitäten zwischen .30 und .62 insgesamt 49% der Varianz klären. Vorerst wurde angenommen, dass die Faktoren unkorreliert sind, die Spezifität der Varianz der Itemsets wurde nicht berücksichtigt: Die Itemsets zur *Deduktion*, *Induktion*, *Korrelation* und *Kausalität* sowie *subjektiven und objektiven Meinungsunterscheidung* laden auf einem gemeinsamen Faktor, der 27% der Varianz klärt. Die Itemsets zur *Zielevaluation* und *Handlungsplanung* laden auf einem gemeinsamen zweiten Faktor, der 21% der Varianz klärt (Anhang, Tabelle 31). Um die geringen Kommunalitäten für die Variablensets *deduktives Schließen* und *objektive und subjektive Meinung unterscheiden* zu berücksichtigen, wurde eine dreifaktorielle Lösung überprüft: Über diese Struktur wird 64% der Varianz geklärt, die Itemsets lassen sich deutlich separieren (Anhang, Tabelle 32).

Die Berücksichtigung der Fehlervarianz zeigt, dass ein gemeinsamer Faktor die Itemsets *deduktives Schließen* und *objektive und subjektive Meinung unterscheiden* jedoch nur wenig erklären kann (Anhang, Tabelle 33).

Damit lassen die Daten den Schluss zu, dass bei einer Aggregation der manifesten Variablen auf Itemsetebene ohne Verletzung der grundlegenden Kriterien zwischen einem Faktor *logisches Schließen im Rahmen der Situationsanalyse* und einem gemeinsamen Faktor für die Itemsets

Zielevaluation und *Handlungsplanung* differenziert werden kann, der als *bewertendes Schließen* im Rahmen *planerischen Vorgehens* bezeichnet werden kann. Beide Faktoren ermöglichen eine Varianzaufklärung von 49%. Eine dreifaktorielle Struktur, die einen Eigenwert knapp unter 1 mitberücksichtigt, zeigt, dass innerhalb der kognitiven Fähigkeiten zur *Situationsanalyse* weiter differenziert werden kann. Dabei bildet das Itemset *objektive und subjektive Meinung unterscheiden* gemeinsam mit *deduktivem Schließen* einen bipolaren Faktor. Die Varianzaufklärung bei drei Faktoren liegt bei 64%. Die exploratorische Überprüfung lässt eine empirisch haltbare Differenzierung in zwei Faktoren, sowie unter knapper Verletzung des Eigenwertkriteriums die Annahme von drei Faktoren zu, wobei *Deduktion* und *Fakten und Meinung unterscheiden* unter Berücksichtigung der Fehlervarianz wenig gemeinsame Varianz teilen.

13.3 Modellüberprüfung

Die Überprüfung der faktoriellen Validität erlaubt Rückschlüsse über die Güte der extrahierten Modelle und Faktoren (Anhang, Tabelle 34): Werden auf Itemsetebene die Modelle der exploratorischen Faktorenanalyse überprüft, so zeigt sich, dass nur in Bezug auf ausgewählte Maße ein einfaktorielles Modell bestätigt werden kann. Sowohl das Verhältnis von χ^2 und df als auch der RMSEA-Index liegen an der noch akzeptablen Grenze, die weiteren Indices weisen darauf hin, dass das Modell noch nicht optimal ist und nicht für die Stichprobe passt.

Mit einem guten Modellfit kann hingegen die zweifaktorielle Lösung bestätigt werden ($\chi^2=10,260$, $df=8$, $p=.247$, $CMIN/df= 1,28$, $RMSEA= .031$, $CFI=.975$, $TLI=.092$, $NFI=.953$). Diese umfassen die Itemsets *deduktives Schließen*, *induktives Schließen*, *korrelatives Schließen* und *subjektive und objektive Information unterscheiden* für den Faktor *Situationsanalyse*. Ein zweiter Faktor *bewertendes und planendes Schließen* umfasst die Itemsets *Handlungsplanung* und *Zielevaluation*.

Um für die weitere Analyse von ausreichend Parametern ausgehen zu können, wurden die Einzelitems herangezogen. Hier zeigte sich, dass die dreifaktorielle Struktur nicht repliziert werden konnte, erst eine Trennung zwischen fünf, bzw. sechs Faktoren kann bessere Modellfits aufweisen. So trennt das fünffaktorielle Modell zwischen den Einzelitems für *deduktives Schließen*, *subjektive und objektive Information unterscheiden*, *induktives Schließen*, *korrelatives Schließen* sowie einem gemeinsamen Faktor *bewertenden und planenden Schließens*, der die Items zur *Handlungsplanung* und *Zielevaluation* umfasst. Das sechsfaktorielle Modell separiert nochmals die Items zur *Handlungsplanung* und *Zielevaluation*. Bei beiden Modellen korrelieren die beiden getrennten Faktoren *induktives Schließen* und *korrelatives Schließen* ($r= .767$, $p\leq .000$) sowie *Handlungsplanung* und *Zielevaluation* ($r= .646$, $p\leq .000$) hoch.

Insgesamt lässt sich daraus folgendes ersehen:

Die Heuristik des Problemlöseprozesses konnte widerspiegelt werden. Die Variablensets zur Erfassung kognitiver Fähigkeiten im Rahmen der *Situationsanalyse* können von Variablensets zur Erfassung kognitiver Fähigkeiten im Rahmen *bewertenden und planenden Schließens* unterschieden werden. *Handlungsplanung* und *Zielevaluation* korrelieren hoch, dies kann darüber erklärt werden, dass beiden Subtests gemeinsame kognitive Anforderungen der Bewertung und Planung zugrunde liegen. Weiterhin spiegeln sich auch empirisch die Gemeinsamkeiten der kognitiven Anforderungen des *induktiven* und *korrelativen Schließens* sowie *des deduktiven Schließens* und *der Unterscheidung von Fakten und Meinung* wider, wobei *Deduktion* und *Fakten und Meinung unterscheiden* wenig gemeinsame Varianz teilen. Theoretisch ist dies darüber zu erklären, dass zwar beiden Itemsets die deduktive Verknüpfung gegebener Informationen zugrunde liegt, beim Unterscheiden zwischen Fakten und Meinung jedoch einschränkende Bedingungen mit zu berücksichtigen sind, das heißt, die subjektive Meinung gegenüber den objektiven Fakten erkannt werden muss. Über konfirmatorische Faktorenanalysen wurde bei einer Modellüberprüfung die Spezifität der Items nochmals deutlich. Auf aggregierter Ebene konnte die zweifaktorielle Lösung der exploratorischen Faktorenanalyse bestätigt werden. Werden jedoch die Einzelitems verglichen, so erfüllt erst ein Modell, das zwischen allen Itemsets differenziert, die statistischen Gütekriterien.

Für die Hypothesen bedeutet dies, dass Hypothese 1 „Über die Subtests lassen sich kognitive Teilkompetenzen differenzieren, die der Heuristik des Problemlöseprozesses zugeordnet werden können“ bestätigt werden kann, mit der Einschränkung, dass die den Subtests *Zielevaluation* und *Handlungsplanung* zugrundeliegenden gemeinsamen kognitiven Anforderungen überwiegen.

Hypothese 2 „Über die Itemsets lassen sich voneinander unterscheidbare Teilkompetenzen des deduktiven, induktiven und korrelatives Schließens sowie der Unterscheidung von Fakten und Meinung differenzieren.“ kann ebenso bestätigt werden. Ebenfalls mit der Einschränkung, dass induktives und korrelatives Schließen hoch korrelieren, das heißt, den Items liegen überwiegend gemeinsame Anforderungen zugrunde.

13.4 Abgrenzung der Testintelligenz

Im Rahmen der Studie wurde Testintelligenz als Intelligenzscreening über 12 Matrizen des Set I des APM (Raven, dt. Bullheller & Häcker, 2005) erhoben. Die Matrizen fokussieren auf kulturfreies analytisch-deduktives Schließen, ihre Lösung erfordert die Fähigkeit der Mustererkennung und Musterfolgeergänzung (Bullheller & Häcker, 2005). Im Folgenden wird geprüft, inwieweit manifeste fächerübergreifende Problemlösekompetenzen und Testintelligenz in Abhängigkeit der zugrundeliegenden Anforderungsstruktur der fächerübergreifenden Problemstellung über einen gemeinsamen Faktor beschrieben werden können.

Eine Überprüfung der Faktorstruktur zeigt, dass unter Einbeziehung der Testintelligenz drei Faktoren extrahiert werden können, die 57% der Gesamtvarianz erklären (Anhang, Tabelle 36). Die drei Faktoren können über die ihnen zugrundeliegenden kognitiven Anforderungen separiert werden: So können *induktives* und *korrelatives Schließen* über einen gemeinsamen Faktor beschrieben werden *Handlungsplanung* und *Zielbewertung* über einen gemeinsamen Faktor *bewertendes und planendes Schließen*. Ein dritter Faktor wird über die Testintelligenz, die Itemsets zum *deduktiven Schließen* sowie die Itemsets zum *Abgleich von Fakten und Meinung* gebildet, wobei der Faktor mit Kommunalitäten $h^2 > .4$ mehr zur Klärung *deduktiven Schließens* sowie der *Testintelligenz* beiträgt als zum *Unterscheiden von Fakten und Meinung* ($h^2 = .4$; Anhang, Tabelle 36). Eine vierfaktorielle Struktur mit einem vierten Faktor mit einem Eigenwert von $\lambda < 1$ und einer Varianzklärung von insgesamt 70% verdeutlicht nochmals die gemeinsame Varianz *deduktiven Schließens* und der *Testintelligenz* (Anhang, Tabelle 37).

Die Berücksichtigung der Fehler- und spezifischen Variablenvarianz bekräftigt nochmals, dass *Testintelligenz* und *deduktives Schließen* mehr Gemeinsamkeiten aufweisen als *Testintelligenz* und *Fakten und Meinung unterscheiden*, zeigt jedoch über die geringen Kommunalitäten auch, dass der gemeinsame Faktor für *Testintelligenz* und *deduktives Schließen* nur wenig Varianz erklärt (Anhang, Tabelle 38).

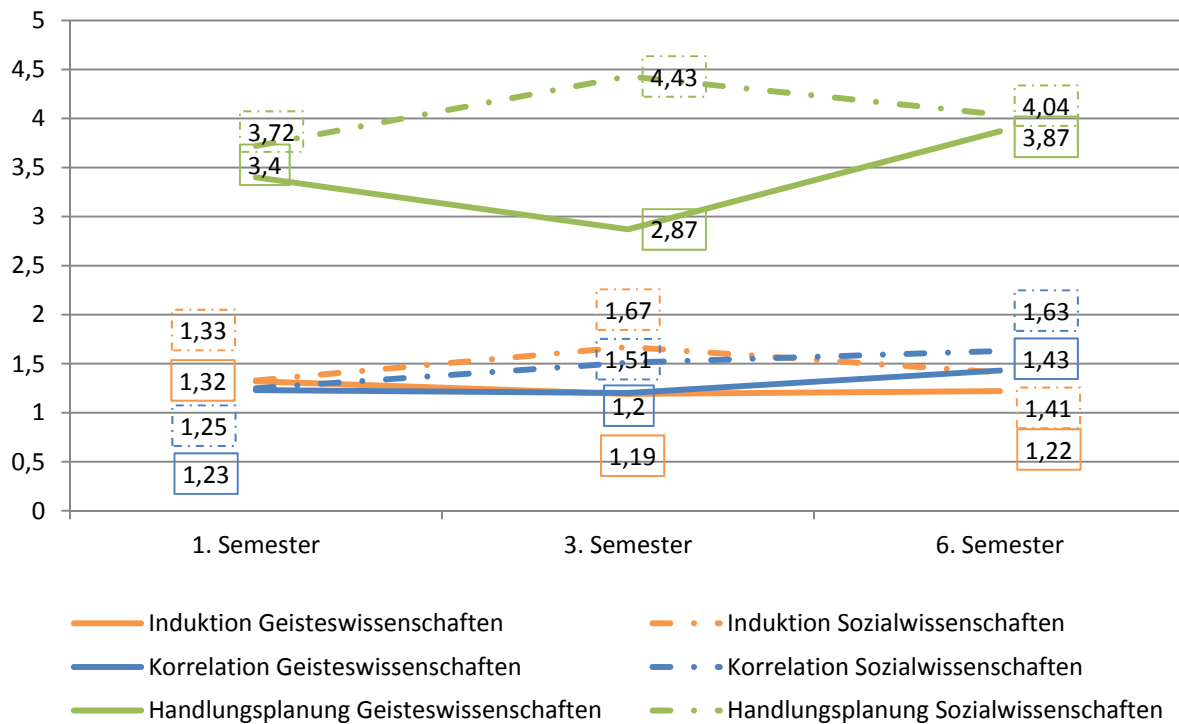
In Bezug auf Hypothese 3 lässt sich dementsprechend festhalten, dass die Testitems zum fächerübergreifenden Problemlösen in Abhängigkeit der kognitiven Anforderung gemeinsame Varianz teilen. Das bedeutet, dass kulturfreie, nicht-sprachliche *Testintelligenz* von *induktivem, korrelativem Schließen* sowie kognitiven Anforderungen der *Zielevaluation* und *Handlungsplanung* unterschieden werden kann, wohingegen die Variablensets zum *deduktiven Schließen* sowie *Fakten und Meinung unterscheiden* in geringem Maß über gemeinsame Anforderungen mit den Testitems des APM beschrieben werden können. Dennoch konnte die Faktorenanalyse deutlich zeigen, dass die Testitems zum fächerübergreifenden Problemlösen, auch zum *deduktiven Schließen*, mehr messen als dekontextualisierte *Testintelligenz*.

13.5 Domänenspezifität fächerübergreifenden Problemlösens

Es wurde bereits dargestellt, dass die Frage nach der Domänenspezifität fächerübergreifender Problemlösekompetenzen bislang ungeklärt ist (Kap. 5.3). Im Rahmen der Studie wurde die Frage insofern überprüft, dass Sozialwissenschaftsstudierende und Geisteswissenschaftsstudierende in ihren fächerübergreifenden Problemlösekompetenzen verglichen wurden. Ein statistisch bedeutsamer Unterschied im Sinne einer besseren Problemlösekompetenz einer Studierendengruppe wäre ein Indiz für die Annahme, dass die Studierenden (im Rahmen ihres Studiums) mit entsprechenden Problemen konfrontiert wurden, bei der Testung einen Transfer erbringen konnten und entsprechend eine bessere Problemlöseleistung zeigen.

Über T-Tests für unabhängige Stichproben wird geprüft, ob sich die Mittelwerte zwischen den Disziplinen Sozialwissenschaften (Schulpsychologie und Journalistik) und Geisteswissenschaften (Anglistik und Europastudien) signifikant unterscheiden. Aufgrund der bereits beschriebenen Verletzung der Normalverteilung werden parametrische und nicht-parametrische Verfahren (Mann-Whitney-U-Test) verwendet, bei Varianzheterogenität entsprechend robuste Verfahren (Welch-Test). Die Ergebnisse zeigen, dass im Rahmen des Subtests *Situationsanalyse* sowie mit Blick auf den Subtest *Handlungsplanung* signifikante Mittelwertunterschiede zu beobachten sind (Anhang, Tabelle 39). Effektstärken mit $d \geq .2$ ($p < .05$) zeigen, dass in kleinem Maß Sozialwissenschaftsstudierende statistisch bedeutsam bessere Leistungen beim *induktiven* und *korrelativen Schließen* sowie bei den Testitems zur *Handlungsplanung* zeigen. Hypothese 4 „Studierende der Sozialwissenschaften und Studierende der Geisteswissenschaften unterscheiden sich im Hinblick auf ihre fächerübergreifende Problemlösekompetenz. Es wird angenommen, dass Studierende der Sozialwissenschaften bessere Leistungen bei der Lösung von Problemen zeigen, die auf induktivem und kausalem Schließen basieren“ kann dementsprechend bestätigt werden. Darüber hinaus weisen die Befunde auf bedeutsam bessere Leistungen der Sozialwissenschaftsstudierenden bei der *Handlungsplanung* im fächerübergreifenden Kontext hin. Die Berücksichtigung des Faktors der Expertise spielt allerdings eine Rolle (Anhang, Tabelle 40). So zeigen sich lediglich im dritten Semester signifikante Unterschiede zwischen den Disziplinen. Zwischen den ersten und sechsten Semestern bestehen keinerlei statistisch bedeutsame Unterschiede. Werden die in Tabelle 40 dargestellten Mittelwerte in einer Grafik nochmals vergleichend abgetragen und die Variablen induktiven und korrelativen Schließens sowie der Handlungsplanung herausgegriffen, so zeigen sich dennoch höhere Leistungen der Sozialwissenschaftler über die drei Semester hinweg.

Grafik 9: Vergleich der Disziplinen und drei Semester bzgl. Induktion, Korrelation und Handlungsplanung



14. Fragenkomplex II: Wirkungszusammenhänge zwischen personbezogenen Variablen und fächerübergreifenden Problemlösekompetenzen

Um fächerübergreifende Problemlösekompetenzen differentialpsychologisch zu beschreiben, wurde nach dem Einfluss personbezogener Kovariaten gefragt. Für die vorliegende Untersuchung wurden die Faktoren Expertise, operationalisiert über die Zahl der Studiensemester, Studieninteresse (FSI, Schiefele, Krapp & Winterler, 1993), Selbstwirksamkeit (adaptierte Skala nach Schwarzer & Jerusalem, 1999) sowie die epistemologischen Überzeugungen (adaptierte Skala nach Urhahne & Hopf, 2004; Transfer: eigenes Design) untersucht (zur Theorie Kap. 5.5).

Um die Wechselwirkungen zwischen personalen Faktoren und fächerübergreifender Problemlösekompetenz zu skizzieren, wurde explorativ eine Zusammenhangsanalyse durchgeführt. Daneben wurden folgende Hypothesen überprüft: Es wird erwartet, dass ...

- ... Studierende mit einer höheren Selbstwirksamkeit fächerübergreifende Probleme besser lösen können als Studierende mit einer geringen Selbstwirksamkeit. (Hypothese 5).
- ... Studierende mit einem höheren Studieninteresse fächerübergreifende Probleme besser lösen können als Studierende mit einem geringeren Studieninteresse. (Hypothese 6).
- ... Studierende mit reflektierteren epistemologischen Überzeugungen eine höhere Problemlösekompetenz haben als Studierende mit weniger reflektierten epistemologischen Überzeugungen (Hypothese 7).

- ... Wechselwirkungen zwischen Selbstwirksamkeit, epistemologischen Überzeugungen und fächerübergreifender Problemlösekompetenz vorliegen. Es wird exploriert, inwieweit eine höhere Selbstwirksamkeit mit reflektierteren epistemologischen Überzeugungen und einer höheren fächerübergreifenden Problemlösekompetenz einhergeht (Forschungsfrage 8).

14.1 Verwendete Verfahren

Für einen ersten Überblick wurden korrelative Zusammenhänge geprüft. Die Darstellung der Zusammenhänge erfolgt über den Korrelationskoeffizient Pearson's r , die Frage nach der Vorhersageverbesserung der fächerübergreifenden Problemlösekompetenz über die Zusammenhangsmaße Eta η , Lambda λ bzw. Cramers V . In einem weiteren Schritt wurde überprüft, inwieweit personbezogene Kovariaten als Prädiktoren der Testleistung herangezogen werden können. Unterschiedsanalysen wurden über T-Tests, der Erklärungsbeitrag über Regressionsanalysen berechnet. Bei Daten auf Basis einer Ratingskala wurden neben verteilungsfreien Verfahren (Nagelkerke, Kox und Snell) ebenfalls parametrische Verfahren durchgeführt (zur Kontroverse Bortz & Döring, 2006). Abschließend überprüfen Strukturgleichungsmodelle die Validität der berichteten Beziehungen. Den Fragebögen lag eine vierstufige Ratingskala von „1 trifft gar nicht zu“ bis „4 trifft genau zu“ zugrunde. Inverse Items wurden umgepolt. Eine höhere Ausprägung entspricht einer reflektierteren Überzeugung (Selbstwirksamkeit, Interesse, epistemologische Überzeugungen). Bei fächerübergreifenden Problemlösekompetenzen wird mit z-transformierter Testleistung gerechnet. Sofern notwendig, werden der besseren Lesbarkeit wegen nicht transformierte Mittelwerte berichtet.

14.2 Zusammenhänge zwischen personalen Variablen

Korrelationsanalysen zeigen in einem ersten Überblick, dass Selbstwirksamkeit, Studieninteresse sowie die Überzeugung, dass Wissen und Wissensentwicklung sowohl flexibel als auch unabhängig von wissenschaftlichen Quellen und Autoritäten ist, in statistisch bedeutsamem, unterschiedlich starkem Zusammenhang stehen ($r=.15-.51$, $p<.01$; Anhang, Tabelle 43; Konvention der Höhe der Korrelation Cohen, 1988). Signifikante mittlere Zusammenhänge zeigen sich ebenso mit der Überzeugung der Transferierbarkeit fachspezifischer Arbeitsmethoden auf fachübergreifende Problemstellungen ($r=.34 - .38$, $p<.01$).

Das bedeutet, im Rahmen der Studie zeigen schwache Zusammenhänge, dass Studierende mit einem höheren Studieninteresse eher von der Wirksamkeit ihres Handelns überzeugt sind und umgekehrt, dass Studierende mit einer höheren Selbstwirksamkeitserwartung interessierter an ihrem Studienfach sind.

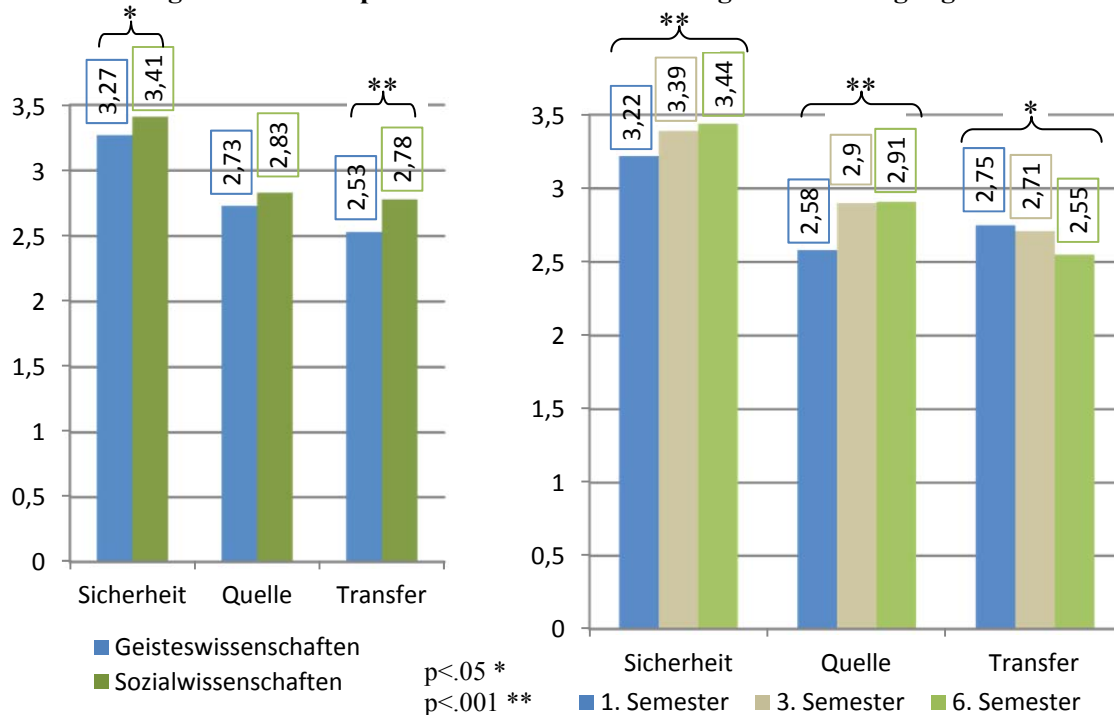
Die Befunde zeigen weiter, dass Studierende mit einer höheren Selbstwirksamkeitserwartung und Studierende mit einem höheren Studieninteresse eher davon überzeugt sind, dass ihre fachspezifisch erlernten Problemlösemethoden auf fächerübergreifende, neue Probleme übertragen werden können.

Das bedeutet, Studierende, die weniger an ihrem Studium interessiert sind, gehen weniger von einer Transfermöglichkeit fachspezifischer Problemlösemethoden aus. Ein korrelativer Zusammenhang konnte weiterhin zwischen Selbstwirksamkeitserwartung und epistemologischen Überzeugungen (Sicherheit und Quelle) gefunden werden. Im Weiteren zeigen die Befunde, dass Studierende mit einer höheren Selbstwirksamkeitserwartung eher davon ausgehen, dass Wissen veränderbar und flexibel ist (Sicherheit des Wissens). Zudem gehen Studierende mit einer höheren Selbstwirksamkeitserwartung davon aus, dass Wissen ein Ergebnis unabhängiger Konstruktionsprozesse ist, das heißt, sie zeigten eine reflektiertere Überzeugung im Hinblick auf die Quelle des Wissens.

Insgesamt zeigen die Befunde starke Zusammenhänge zwischen den epistemologischen Überzeugungen: Studierende, die eher von der Absolutheit des Wissens ausgehen sind auch davon überzeugt, dass Wissen von Autoritäten weitergegeben wird. Andererseits bedeutet dies, dass die Überzeugung im Hinblick auf die Veränderbarkeit von Wissen (Sicherheit) mit der Überzeugung in Zusammenhang steht, dass Wissen auf der Fähigkeit des Lernenden aufbaut und in der Interaktion und Auseinandersetzung mit anderen weiterentwickelt werden kann (Quelle). Diese Beziehungen konnten im Rahmen eines Strukturgleichungsmodells mit einem guten Modellfit überprüft werden ($\chi^2= 12,10$; $df= 8$, $p=.147$, $CMIN/df=1,513$; $RMSEA=.040$, $CFI= .975$, Kap. 14.5).

Was lässt sich bei einem Vergleich der drei Semester sowie der beiden Disziplinen feststellen? Erstsemester sind signifikant überzeugter als Sechstsemester, dass fachspezifische Problemlösemethoden auf fächerübergreifende Problemstellungen übertragen werden können. Demgegenüber differenzieren sich die epistemologischen Überzeugungen bezüglich Wissen und Wissensentwicklung mit zunehmendem Semester aus. Das bedeutet, je höher das Semester, umso weniger überzeugt sind die Studierenden vom Transfer fachspezifischer Problemlösemethoden auf neue und nicht-fachgebundene Probleme, ihre epistemologischen Überzeugungen sind reflektierter (Anhang, Tab. 41).

Ein Vergleich der Studiendisziplinen zeigt einen signifikanten Unterschied zwischen den Disziplinen im Hinblick auf die Sicherheit des Wissens sowie den Transfer der Problemlösemethoden: Die Befunde zeigen, dass Sozialwissenschaftsstudierende überzeugter sind, fachspezifische Problemlösemethoden auf fächerübergreifende Problemstellungen anwenden zu können als Geisteswissenschaftsstudierende. Ebenso sind sie überzeugter, dass Wissen veränderbar und flexibel ist (Anhang, Tab. 42).

Grafik 10: Vergleich der Disziplinen und Semester in Bezug auf Überzeugungen

In Bezug auf die Überzeugung, dass das eigene Handeln wirksam ist, sowie in Bezug auf das Interesse am Fachstudium zeigt sich, dass sich die Disziplinen statistisch bedeutsam unterscheiden. Die Studierenden der Sozialwissenschaften sind sowohl interessierter als auch eher von der Wirksamkeit ihres Handelns überzeugt (Selbstwirksamkeit: $T=-2,24$, $df=259$, $d=3$, $p<.05$; Interesse: $T=-2,07$, $df=259$, $d=3$, $p<.05$). Im Weiteren zeigt sich, dass die Erstsemester signifikant interessierter an ihrem Fachstudium sind als Dritt- und Sechstsemester ($F=4,51$, $df=2/258$, $p<.05$, $d_{1.-3.Sem.}=4$, $d_{1.-6.Sem.}=4$). Dritt- und Sechstsemester unterscheiden sich nicht. Im Hinblick auf die Selbstwirksamkeit unterscheiden sich die Semester ebensowenig ($F=.56$, $df=2/259$, $p>.05$). Für einen deskriptiven Überblick siehe Anhang, Tabelle 44.

Bei der Frage, ein inwieweit geschlechtsspezifischer Bias festgestellt werden kann, zeigt sich, dass Testfairnis vorliegt, allerdings zeigen sich signifikante Unterschiede zwischen Männern und Frauen im Hinblick auf Selbstwirksamkeit und Studieninteresse. Dabei muss die Ungleichverteilung zwischen den Geschlechtern berücksichtigt werden:

Tab. 3: Geschlechtsspezifische Unterschiede

	\bar{x} (SD)		d $\frac{\mu_1 - \mu_2}{\delta}$	T-Test*
	Männer	Frauen		
Gesamttest	11,92 (3,24)	11,31 (3,02)	-	$T=-,905$ ($df=255$) $p=.36$
Situationsanalyse	5,07 (1,42)	4,67 (1,61)	-	$T=-1,174$ ($df=255$) $p=.09$
Zielevaluation	2,74 (,81)	2,89 (,93)	-	$T=-1,121$ ($df=255$) $p=.26$

Handlungsplanung	4,12 (2,51)	3,76 (2,01)	-	T=-1,129 (df=255) p=.26
Transfer	2,73 (,52)	2,65 (,54)	-	T=-,974 (df=254) p=.33
Sicherheit	3,37 (,43)	3,34 (,40)	-	T=-,484 (df=255) p=.63
Quelle	2,85 (,59)	2,76 (,54)	-	T=-1,136 (df=255) p=.26
Interesse	2,89 (,60)	3,15 (,53)	d=.4	T=-3,71 (df=254) p=.00
Selbstwirksamkeit	3,11 (,40)	2,99 (,40)	d=.3	T=-1,973 (df=254) p=.05

*nicht-parametrische Tests bestätigen die Signifikanzen auf gleichem Niveau (Mann-Whitney-U), mit einer Signifikanz von $p=.07$ nicht mehr bei Selbstwirksamkeit.

Frauen $N=200$, Männer $N=57$

Mit der Analyse von Prädiktorvariablen lassen sich über multiple Regressionsanalysen (Prozedur: schrittweise Aufnahme) die Wechselwirkungen zwischen den signifikant in Zusammenhang stehenden Kovariaten nochmals näher beschreiben. Vor allem stellt sich die Frage, inwieweit bei Kenntnis des Studieninteresses, der Selbstwirksamkeitserwartung sowie einer Komponente der beiden erhobenen epistemologischen Überzeugungen deren Ausprägung vorhergesagt werden kann. Regressionsanalytisch zeigt sich, dass sowohl das Interesse der Studierenden als auch die Selbstwirksamkeitserwartung einen Erklärungsbeitrag für die Überzeugung bezüglich der Transferierbarkeit fachspezifischer Problemlösemethoden auf fächerübergreifende Probleme leisten ($F=34,46$; $df = 2/258$; $R^2_{\text{korr}}=.21$, $\beta_{\text{Studieninteresse}}=.32$, $\beta_{\text{Selbstwirksamkeit}}=.27$, $p<.00$). Das heißt, die Höhe des Studieninteresses und die Erwartung, dass das eigene Handeln wirksam ist, lässt bei den befragten Studierenden einen Schluss zu auf ihre Überzeugung von der Transferierbarkeit fachspezifisch erlernter Problemlösemethoden auf neuartige, fächerübergreifende Probleme.

Die erklärungs mächtigste Variable in Bezug auf die Einschätzung der Flexibilität des Wissens (Sicherheit) ist die Überzeugung, dass Wissen auch über eigene Lernprozesse entwickelt und unabhängig von Autoritäten erworben werden kann (Quelle), in geringerem Maß kann auch die Selbstwirksamkeitserwartung als Prädiktor herangezogen werden ($F= 50,31$; $df= 2/258$; $R^2_{\text{korr}}= .28$; $\beta_{\text{Quelle}}=.48$, $\beta_{\text{Selbstwirksamkeit}}=.16$, $p<.00$). Die Überzeugung, dass Wissen auf der Fähigkeit des Lernenden aufbaut (Quelle) lässt sich darüber vorhersagen, für wie flexibel bzw. absolut Studierende Wissen halten ($F= 88,48$; $df= 1/259$; $R^2_{\text{korr}}= .25$, $\beta_{\text{Sicherheit}}=.51$, $p<.000$). Die ordinale Regression bestätigt diese Vorhersagemodelle mit akzeptablen bis guten Werten (Cox und Snell sowie Nagelkerke $>.2$, Backhaus, Anhang, Tabellen 45-48).

14.3 Differentialpsychologische Faktoren als Kovariaten der Problemlöseleistung

Wie sehr beeinflussen motivationale und metakognitive Variablen fächerübergreifende Problemlösekompetenzen? Korrelativ zeigen sich schwache Zusammenhänge zwischen den epistemologischen Überzeugungen (Sicherheit, Quelle) und dem Gesamtergebnis (Sicherheit

$r=.20$, $p<.001$; Quelle $r=.14$, $p<.001$) sowie der Leistung im Subtest Handlungsplanung (Sicherheit $r=.16$, $p<.05$; Quelle $r=.12$, $p<.05$) (Anhang, Tabelle 49). Auf Ebene der Variablensets zeigen sich Zusammenhänge zwischen der Überzeugung von der Wirksamkeit des eigenen Handelns und der Leistung in der Unterscheidung zwischen Fakten und Meinung ($r=.18$, $p<.05$). Weitere Zusammenhänge zwischen Selbstwirksamkeit, bzw. Studieninteresse und fächerübergreifenden Problemlösekompetenzen lassen sich nicht beobachten. Die negativen Zusammenhänge zwischen epistemologischen Überzeugungen und deduktivem Schließen lassen sich über den Einfluss des Semesters erklären: Sechstsemester haben reflektiertere Überzeugungen, lösen jedoch die Testitems zum deduktiven Schließen schlechter (vgl. Kap.15.2).

Wird in einem nächsten Schritt danach gefragt, inwieweit die Kenntnis einer personbezogenen Kovariate die Vorhersage der Testleistung verbessert (PRE-Maße, *Proportional Reduction in Error*), zeigt sich, dass über die Kenntnis der epistemologischen Überzeugungen das Testergebnis in den Subtests Situationsanalyse und Handlungsplanung statistisch bedeutsam besser vorhergesagt werden kann. Studieninteresse, Selbstwirksamkeit sowie die Überzeugung von der Transferierbarkeit fachspezifischer Problemlösemethoden auf fächerübergreifende Probleme haben keinen Einfluss (Anhang, Tabelle 50).

Vor diesem Hintergrund stellt sich die Frage, welche personbezogenen Faktoren als Prädiktoren einen Erklärungsbeitrag für die fächerübergreifenden Problemlösekompetenzen leisten. Die Befunde zeigen, dass Studierende, die von der Veränderbarkeit des Wissens überzeugt sind, bessere Problemlöseleistungen auf Gesamttestebene zeigen (Anhang, Tabelle 51). Mit β -Gewichten zwischen $\beta=.15$ - $.16$ ($p<.05$) lässt sich die Überzeugung von der Veränderlichkeit des Wissens als Prädiktor für die Subtests Situationsanalyse und Handlungsplanung heranziehen (Anhang, Tabelle 52-53). Eine Analyse auf Subtestebene zeigt kleine, aber konsistente Effekte bei Testitems, die auf induktivem Schließen ($R^2=.03$, $T=2,87$, $\beta=.18$, $p<.05$) und insbesondere auf korrelativem Schließen ($R^2=.09$, $T=4,970$, $\beta=.30$, $p<.001$) basieren; ebenso kleine Effekte wie bei induktivem Schließen, allerdings ebenfalls signifikant ist der Einfluss auf die Fähigkeit zur Handlungsplanung ($R^2=.03$, $T=2,56$, $\beta=.16$, $p<.05$). Ein Vergleich zwischen weniger reflektierten und reflektierteren Studierenden verdeutlicht den Effekt der Überzeugung von der Veränderlichkeit des Wissens auf induktives ($d=.4$, $p<.05$) und korrelatives Schließen ($d=.7$, $p<.00$) (Mediansplit, Anhang, Tabelle 51). Das bedeutet, Studierende, die ihr fachspezifisch erlerntes Wissen als veränderlich ansehen, lösen diejenigen Probleme signifikant besser, die den Umgang mit Unsicherheiten sowie das Erkennen einschränkender Bedingungen erfordern.

14.4 Testintelligenz und Expertise

Inwieweit kann Testintelligenz zur Erklärung der Testleistung herangezogen werden? Hier zeigt sich, dass Intelligenz im Hinblick auf die Gesamttestleistung ($\beta=.15$, $p<.05$) sowie mit Blick auf die

Testleistung bei den Subtests Situationsanalyse ($\beta=.13$, $p<.05$) einen Teil der Subtestleistung erklären kann (Anhang, Tabelle 55-56). Ausgehend von bisherigen Befunden zum Zusammenhang zwischen Intelligenz und Problemlösen (für einen Überblick Kap.5.5) sowie in Anlehnung an die Elshout-Raaheim-Hypothese (Funke, 2006), die das Zusammenwirken von Intelligenz und Vorwissen beim Problemlösen herausstellt, wurde danach gefragt, inwieweit sich der Zusammenhang zwischen Testintelligenz und Testleistung in Abhängigkeit der Problemlösefähigkeit verändert. Hierzu wurde die Gesamtstichprobe in Abhängigkeit der Gesamtestleistung in zwei ähnlich zusammengesetzt Teilstichproben gesplittet, die „schwächeren Problemlöser“ und die „stärkeren Problemlöser“. Bei der Teilstichprobe der schwächeren sind im Gegensatz zur Teilstichprobe der stärkeren Problemlöser schwache korrelative Zusammenhänge zwischen Testintelligenz und dem Gesamtestergebnis, bzw. dem Subtest Handlungsplanung festzustellen (Anhang, Tabelle 57). Weiterhin sind – im Gegensatz zur Teilstichprobe der stärkeren Problemlöser ($r_{\text{SW-Gesamtest}}=-.06$, $p>.05$; $r_{\text{SW-Handlungsplanung}}=.02$, $p>.05$), auch schwache, jedoch signifikante Zusammenhänge zwischen dem Gesamtestergebnis ($r=.18$, $p<.05$) bzw. dem Subtest Handlungsplanung ($r=.17$, $p<.05$) und Selbstwirksamkeit zu beobachten. Bei der Frage der Vorhersage der Testleistung über die Ausprägung der epistemologischen Überzeugungen sowie der Testintelligenz sind weiterhin höhere β -Gewichte festzustellen sowohl in Bezug auf die Überzeugung von der Sicherheit des Wissens (Gesamttest $\beta=.23$, $p<.05$), als auch in Bezug auf den Faktor der Testintelligenz (Gesamttest $\beta=.19$, Handlungsplanung $\beta=.22$, beide $p<.05$) (Anhang, Tabellen 58-60). In Bezug auf die Testleistung im korrelativen Schließen zeigt sich ein sehr viel bedeutsamerer Einfluss der Überzeugung von der Flexibilität des Wissens ($R^2=.15$, $T=4,78$, $\beta=.38$, $p<.000$). Aufgrund der Problematik der Berechnungen auf Basis des Mediansplits sollen die Befunde an dieser Stelle lediglich als Ansatzpunkt für die weiterführende Frage dienen, inwieweit sich in Abhängigkeit der fächerübergreifenden Problemlösekompetenz die Relevanz der Ausprägung personbezogener Kovariaten wie der Selbstwirksamkeit, epistemologischer *beliefs* sowie der Testintelligenz verändert.

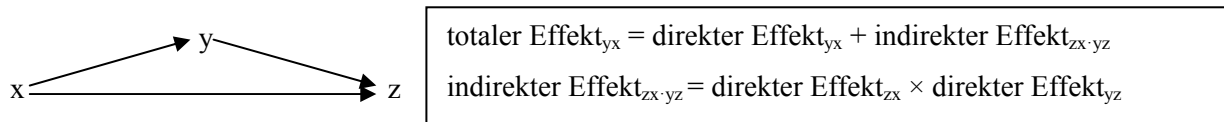
14.5 Zusammenfassung der Wechselwirkungen und Skizze eines Strukturmodells

Nach der explorativen Darstellung der Zusammenhänge gibt eine abschließende Modellierung direkter und indirekter Zusammenhänge im Rahmen eines Strukturmodells einen zusammenfassenden Überblick und anhand der Modellfits Aufschluss über die Güte der Zusammenhänge. Die Pfadanalysen visualisieren die Stärke der Zusammenhänge. Die Kriterien zur Beurteilung der über das Maximum-Likelihood-Prinzip geschätzten Modelle wurden bereits vorgestellt (Kap.13.1). In der nachstehenden Grafik sind lediglich diejenigen Faktoren enthalten, die einen statistisch bedeutsamen Beitrag zum Modell leisten. Die Pfadgewichte sind als standardisierte Schätzungen angegeben. Auf eine Abbildung des Messmodells wurde aus Gründen der Übersichtlichkeit verzichtet, eine

Überprüfung der Güte des Messmodells erfolgte bereits im Rahmen der Pilotierung, eine Auseinandersetzung mit der Binnenstruktur des Konstruktes in Fragenkomplex I.

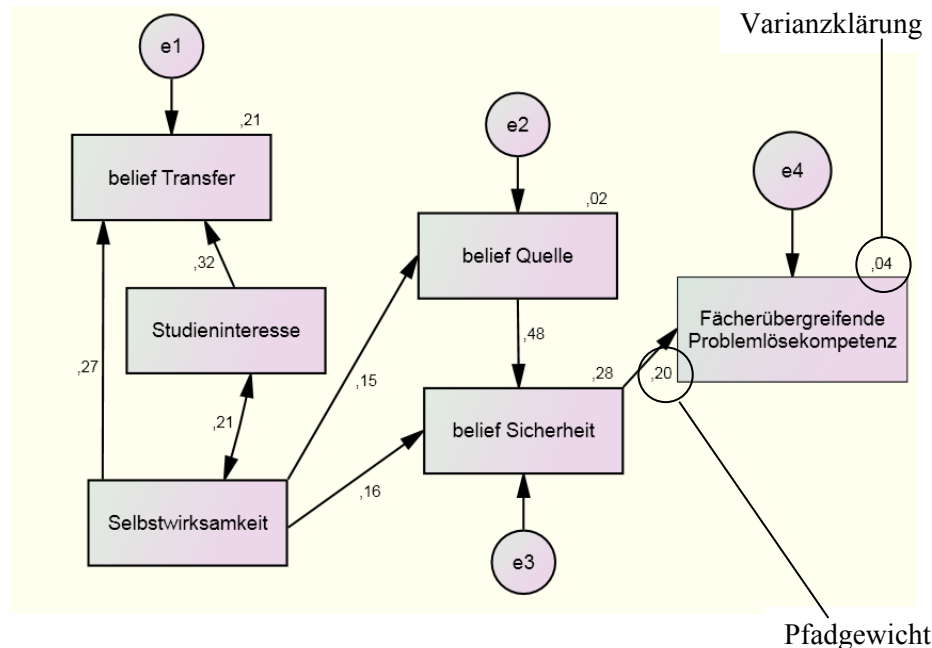
Das Modell zeigt, dass fächerübergreifende Problemlösekompetenzen gemessen über das Gesamtergebnis über die untersuchten Kovariaten direkt und indirekt beschrieben werden können. Eine Tabelle der Effekte findet sich im Anhang und folgt für die Berechnung folgenden Gleichungen (Anhang, Tabelle 61-62):

Abb. 5. Totale und indirekte Effekte



Es wurden auf Basis der Befunde unterschiedliche ‚nested models‘ überprüft, die aus Gründen der Übersichtlichkeit nicht aufgeführt sind. Die Modellfits zeigen, dass nachstehendes Strukturmodell die Beziehungen für die Stichprobe sehr gut beschreibt ($\chi^2 = 12,10$; $df = 8$, $p = .147$, $CMIN/df = 1,51$, $RMSEA = .040$, $CFI = .975$): Sowohl Selbstwirksamkeit als auch Studieninteresse nehmen Einfluss darauf, inwieweit fachspezifische Problemlösemethoden als übertragbar auf fachübergreifende Problemstellungen angesehen werden. Das heißt, je interessierter Studierende an ihrem Fachstudium und je überzeugter sie von ihrer eigenen Handlungsfähigkeit sind, umso eher gehen sie davon aus, dass ihre fachspezifisch erworbenen Problemlösemethoden auch bei der Bearbeitung und Lösung neuer und fächerübergreifender Probleme hilfreich sind. Studieninteresse und Selbstwirksamkeit selbst stehen in bidirektionaler Beziehung, das bedeutet, je interessierter am Fachstudium, umso überzeugter sind Studierende von der Wirksamkeit ihres Handelns und umgekehrt. Mit einer etwas geringeren Pfadstärke nimmt Selbstwirksamkeit statistisch bedeutsam Einfluss auf die epistemologischen Überzeugungen. Das heißt, eine reflektierte Überzeugung in Bezug auf Wissen und die Wissensentwicklung zeigt sich in vorliegender Studie mit einer Pfadstärke von $r = .2$ abhängig von der Einschätzung der Wirksamkeit des eigenen Handelns. Als Prädiktor fächerübergreifender Problemlösekompetenz kann die Überzeugung bezüglich der Flexibilität des Wissens bestätigt werden.

Grafik 11: Strukturmodell zu Wechselwirkungen und Prädikation fächerübergreifender Problemlösekompetenzen



Die Eignung des Strukturmodells wurde über die Überprüfung an einer aus der Gesamtstichprobe zufällig gezogenen Teilstichprobe als ‚validation sample‘ nochmals überprüft (Diamantopoulos & Siguaw, 2000). Die Modellfits bleiben bei einem Wert von CFI >.9, der Chi-Quadratstest bleibt nicht signifikant.

Was kann abschließend festgehalten werden?

Die Ergebnisse zeigen, dass Hypothese 5 „Es wird erwartet, dass Studierende mit einer höheren Selbstwirksamkeit fächerübergreifende Probleme besser lösen können als Studierende mit einer geringen Selbstwirksamkeit.“ nur eingeschränkt angenommen werden kann und der Einfluss der Selbstwirksamkeit auf fächerübergreifendes Problemlösen nicht direkt beschrieben werden kann. Inwieweit Selbstwirksamkeit insbesondere bei schwächeren Problemlösern relevant wird muss noch weiter überprüft werden. Die Befunde zeigen, dass Selbstwirksamkeit in erster Linie indirekt über die metakognitiv verorteten epistemologischen Überzeugungen – insbesondere die Überzeugung von der Flexibilität des Wissens, Einfluss auf fächerübergreifende Problemlösekompetenz nimmt. Des Weiteren nimmt Selbstwirksamkeit Einfluss auf die Überzeugung bezüglich des Transfers fachspezifischer Problemlösemethoden auf fachübergreifende Problemstellungen, sofern diese Komponente zu den metakognitiv verorteten *beliefs* zugeordnet wird, auch wenn, wie die Befunde zeigen, die epistemologischen Überzeugungen von der Transferüberzeugung unabhängig zu sehen sind. Die Ergebnisse zeigen, dass Selbstwirksamkeit keinen bedeutsamen Beitrag zur Erklärung der Testleistung bzw. der Leistung in den Subtests leistet. Selbstwirksamkeit scheint damit ein relevanter, jedoch indirekt wirksamer Faktor zu sein.

Hypothese 6, „Es wird erwartet, dass Studierende mit einem höheren Studieninteresse fächerübergreifende Probleme besser lösen können als Studierende mit einem geringeren Studieninteresse“ muss abgelehnt werden. Die Befunde legen keinen Zusammenhang mit Studieninteresse und fächerübergreifenden Problemlösekompetenzen nahe. Im Rahmen der Diskussion der Befunde wird genauer dargestellt, inwieweit dieses Ergebnis darüber erklärt werden kann, dass das im Rahmen der Studie verwendete Erhebungsinstrument, bzw. die Operationalisierung des Konstrukts *Studieninteresse* zu unspezifisch ist. Weiterhin bleibt danach zu fragen, inwieweit die fächerübergreifende Problemlösekompetenz gegebenenfalls weniger mit Fachstudieninteresse in Zusammenhang steht, als vielmehr mit einem Interesse an Problemstellungen, die über das Studium hinausgehen.

Hypothese 7 „Es wird erwartet, dass Studierende mit reflektierteren epistemologischen Überzeugungen eine höhere Problemlösekompetenz haben als Studierende mit weniger reflektierten epistemologischen Überzeugungen“ kann angenommen werden. Dabei konnte gezeigt werden, dass insbesondere die Überzeugung von der Sicherheit des Wissens fächerübergreifende Problemlösekompetenz beeinflusst, wobei die Leistung reflektierter Studierender in denjenigen Problemlöseaufgaben höher waren, die die Berücksichtigung einschränkender Bedingungen und die Vermeidung übergeneralisierender sowie kausaler Zusammenhänge erforderten.

Die explorativ gefasste Forschungsfrage 8, „inwieweit Wechselwirkungen zwischen Selbstwirksamkeit, epistemologischen Überzeugungen und fächerübergreifender Problemlösekompetenz vorliegen [...] [und] inwieweit eine höhere Selbstwirksamkeit mit reflektierteren epistemologischen Überzeugungen und einer höheren fächerübergreifenden Problemlösekompetenz einhergeht“ konnte im Rahmen des Strukturgleichungsmodells mit einem sehr guten Modellfit beschrieben werden.

Die Befunde zeigen, dass Expertise im Sinne des Studienseesters insofern zwischen Studierenden unterscheidet, dass die epistemologischen Überzeugungen mit zunehmendem Semester differenzierter werden und darüber auch die Kompetenz, fächerübergreifende Probleme lösen zu können, zunimmt. Der Einfluss des Semesters auf die fächerübergreifende Problemlösekompetenzen, bzw. ihre Teilkompetenzen wird im folgenden Kapitel dezidiert beschrieben. Im Hinblick auf die Testfairnis kann kein geschlechtsspezifischer *Bias* beobachtet werden. Männer und Frauen unterscheiden sich allerdings im Hinblick auf das Interesse am Studienfach sowie in geringem Maße mit Blick auf die Selbstwirksamkeit und entsprechen damit auch den dargestellten Befunden (vgl. Teil II). Die Befunde sind vor dem Hintergrund der Ungleichverteilung der Geschlechter jedoch nur eingeschränkt aussagekräftig.

15. Fragenkomplex III: Fächerübergreifende Problemlösekompetenzen bei Studierenden

Im Folgenden werden die über das Testinstrument gemessenen fächerübergreifenden Problemlösekompetenzen der Studierenden beschrieben. Insbesondere werden Unterschiede zwischen Semestern und Disziplinen beleuchtet. Die bisher berichteten Befunde zeigen, dass Semester und Studiendomäne Einfluss auf die epistemologischen Überzeugungen nehmen. Diese wiederum konnten als Prädiktoren fächerübergreifender Problemlösekompetenz herausgearbeitet werden. Ausgehend davon wird im Folgenden danach gefragt, inwieweit insbesondere epistemologische Überzeugungen Einfluss auf Teilkompetenzen fächerübergreifenden Problemlösens nehmen und weiterhin in Zusammenhang mit Kovariaten der Studierenerfahrung – Semester und Studiendisziplin – stehen. Abschließend werden die für die Gesamtttestleistung relevanten kognitiven Teilkompetenzen in ihrem Zusammenhang beschrieben. Auch hier wird die Güte der Beziehungen im Rahmen eines Strukturmodells überprüft. Dabei ist zu beachten, dass die Modellprüfung nicht sensitiv für die Wahrscheinlichkeit einer bestimmten Sequenz von Zuständen ist (Funke & Frensch, 2006). Vielmehr wird über das Strukturmodell kognitiver Kompetenzen die Wahrscheinlichkeit einer bestimmten Kombination kognitiver Zustände zur erfolgreichen Bearbeitung fächerübergreifender Problemlöseaufgaben überprüft.

Im Rahmen des Fragenkomplexes werden die Kompetenzen und Zusammenhänge explorativ analysiert. Darüber hinaus sind auf Basis bisheriger Befunde folgende Hypothesen leitend:

Es wird erwartet, dass ...

- ... Expertise Einfluss nimmt und sich Studierende eines höheren Semesters und Studierende eines niedrigeren Semesters im Hinblick auf ihre Problemlösekompetenz unterscheiden. Es wird angenommen, dass Studierende eines höheren Semesters eine höhere Problemlösekompetenz haben als Studierende eines niedrigeren Semesters (Hypothese 9).
- ... diese Unterschiede vor allem bei schwierigen Aufgaben deutlich werden (Hypothese 10).
- ... Unterschiede zwischen den Disziplinen beobachtbar sind. Es wird angenommen, dass Sozialwissenschaftsstudierende [...] diejenigen Problemstellungen besser lösen können, die Kompetenzen des induktiven und korrelativen Schließens erfordern (Hypothese 11).
- ... kognitive Teilkomponenten fächerübergreifenden Problemlösens, die über ähnliche kognitive Anforderungen beschrieben werden können, auch empirisch in Zusammenhang stehen (Hypothese 12).
- ... kognitive Kompetenzen zur Analyse von Problemsituationen (Subtest Situationsanalyse) in Zusammenhang stehen mit Teilkompetenzen der Bewertung und Handlungsplanung (Subtests Zielevaluation und Handlungsplanung) (Hypothese 13).
- ... Studierende, die die Testitems des Subtests Situationsanalyse besser lösen auch die Testitems der Subtests Zielevaluation und Handlungsplanung besser lösen (Hypothese 14).

15.1 Verwendete Verfahren

Es werden sowohl deskriptive Analysen als auch Unterschiedsanalysen dargestellt. Aufgrund der Verteilung der Daten werden auch im Folgenden parametrische und nicht-parametrische Verfahren verwendet, bei Varianz- und Regressionsanalysen werden die robusteren Kennwerte gewählt (Welch-Test, Bonferroni-Korrektur, Wilks Lambda, Pillais-Spur). Für die Strukturmodelle gelten die bereits berichteten Modellgütekriterien. Weiterhin werden Extremgruppen verglichen, um über den Vergleich von Subgruppen Kontrastierungen klarer zu konturieren. Dabei ist zu berücksichtigen, dass die Befunde auf Basis der Extremgruppenvergleiche Ergebnisse für eine spezifische Stichprobe darstellen und für eine ausgewählte Variablenkonstellation Geltung besitzen, die nicht der Realstichprobe entspricht. Extremgruppenvergleiche können bestimmten Variablen und Variablenkonstellationen mehr Bedeutung zukommen lassen.

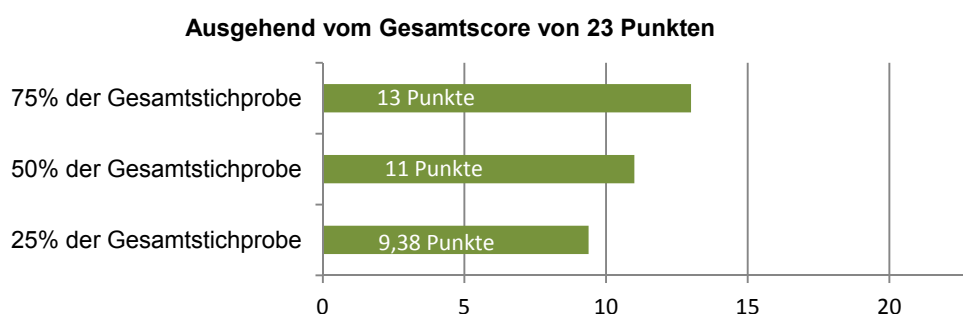
15.2 Fächerübergreifendes Problemlösen – eine Frage der Expertise?

Für einen ersten Einblick werden die Testleistungen der Gesamtstichprobe dargestellt:

Die Schwierigkeitsanalyse im Rahmen der Pilotierung zeigte, dass insbesondere die Testitems zum deduktiven Schließen sowie die Items zur Handlungsplanung besonders schwierig sind (Kap. 10.3). Weiterhin fällt auf, dass die Varianz der Testleistung auf Itemebene geringer ist als auf Itemsetebene. Das bedeutet, die Streuung bei der korrekten Bearbeitung eines Items ist geringer als bei der korrekten Beantwortung beider Items. Dementsprechend werden im Folgenden die Befunde auch auf Einzelitemebene betrachtet und die Einzelergebnisse mit Blick auf die ihnen zugrundeliegende Konstruktion verglichen. Es wird danach gefragt, inwieweit Antwortmuster zu erkennen sind und welche kompetenz- oder itemkonstruktionsbezogenen Gründe bei der Itemlösung eine Rolle spielen könnten.

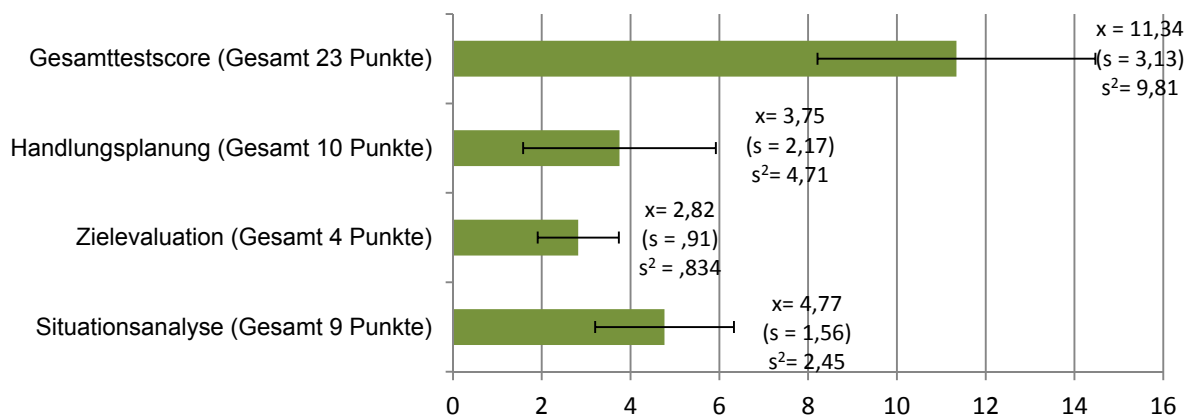
Zur Einordnung der Unterschiede beim fächerübergreifenden Problemlösen stellt die nachstehende Grafik zur prozentualen Verteilung des Gesamtttestscores sowie der Subtestleistungen der Gesamtstichprobe einen Referenzrahmen dar.

Grafik 12: Verteilung der Gesamtttestleistung



Die Grafiken 12 und 13 zeigen, dass bei einem Gesamtscore von 23 Punkten niemand die volle Punktzahl erreichte, im Mittel erreichten die Studierenden insgesamt 11,34 Punkte ($SD=3,13$; $Min=2,0$; $Max=21,0$). Der Blick auf die Quartile verdeutlicht, dass ein Viertel der Studierenden 9 Punkte und weniger erreichte (Q1), die Hälfte der Studierenden 11 Punkte und weniger (Q2), Dreiviertel der Studierenden erreichte 13 Punkte und weniger (Q3) (Grafik 6). Im Mittel erreichten die Studierenden im Subtest Situationsanalyse $\bar{x}=4,77$ ($SD=1,56$) von 9 Punkten, im Subtest Zielevaluation $\bar{x}=2,82$ ($SD=,91$) von 4 Punkten und im Subtest Handlungsplanung $\bar{x}=3,75$ ($SD=1,56$) von 10 Punkten (Grafik 7). Der Blick auf die Varianz zeigt, dass im Subtest Zielevaluation die Testleistung am geringsten streut, die Testleistung im Subtest Handlungsplanung sehr heterogen ist.

Grafik 13: Verteilung und Streuung der Subtestleistung (N=286)



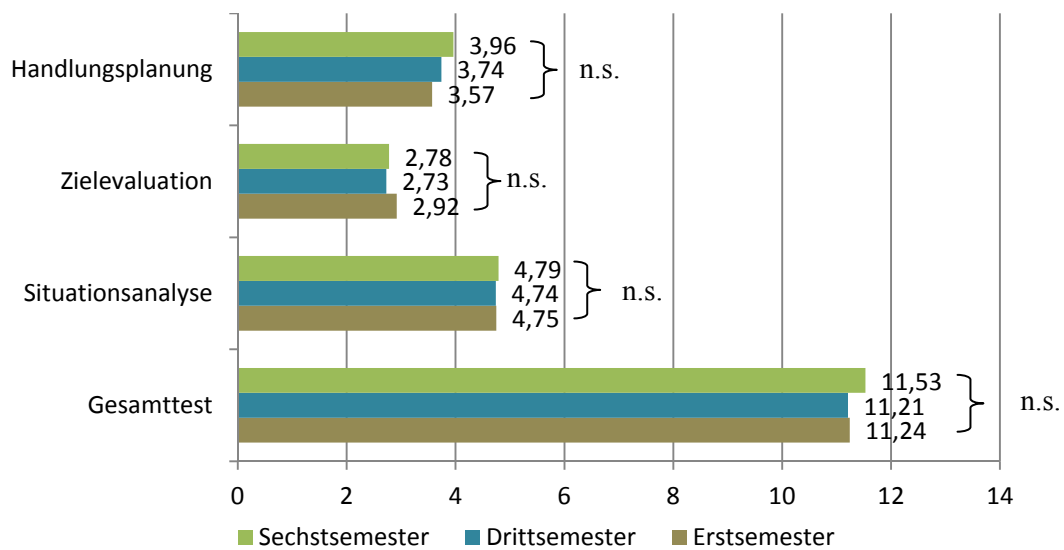
Den folgenden Berechnungen liegen z-transformierte Testergebnisse zugrunde. Zur besseren Lesbarkeit werden begleitend auch Mittelwertunterschiede auf Basis der nicht transformierten Bepunktung der Items, Itemsets oder Subtests genannt.

Der Einfluss des Semesters

Insgesamt wurden N=114 Studierende des ersten Semesters, N=70 Studierende des dritten Semesters sowie N=102 Studierende des sechsten Semesters untersucht. Ein erster Überblick zeigt, dass die nicht transformierte mittlere Testleistung auf Subtest- sowie Gesamtestebene zwischen den Semestern nicht differiert (Mittelwert und Varianzen Anhang, Tabelle 66).

Grafik 14: Überblick über die Testleistung der drei Semester

(arithmet. Mittel, nicht-transformierte Testleistung, N=286)



Ein Vergleich der Subtestleistung zeigt keine statistisch bedeutsamen Unterschiede zwischen den Semestern (Anhang, Tabelle 67). Dies bedeutet, dass Sechstsemesterstudierende keine bedeutsam bessere Testleistung zeigten als Erstsemester. Wird die Streuung betrachtet, so lässt sich feststellen, dass mit fortschreitendem Semester die Testleistung auf Ebene der Subtests etwas breiter streut, das heißt, die Testleistung der Studierenden wird heterogener, jedoch nicht bedeutsam.

Auf Ebene der Variablensets zeigt sich ein ähnliches Bild. Dritt- und Sechstsemester lösen die Testitems zum *Unterscheiden zwischen Fakten und Meinung* und *induktiven Schließen* ebenso wie Erstsemester. Signifikante Unterschiede zwischen den Semestern lassen sich beim *deduktiven* und *korrelativen Schließen* feststellen. Im Hinblick auf korrelatives Schließen zeigt sich, dass die Streuung abnimmt, die Testleistung wird mit zunehmendem Semester homogener (Anhang, Tabellen 66-71).

Zusammenhangsmaße (η) zeigen, dass das Semester für die vorliegende Stichprobe 18% der Testleistung bei deduktiv und 19% bei korrelativ ausgerichteten Testaufgaben erklärt (Anhang, Tabelle 69). Werden erstes und sechstes Semester verglichen, so verändert sich die statistische Bedeutsamkeit der Differenzen beim *induktiven Schließen*, *Fakten und Meinung unterscheiden*, *Zielevaluation* und *Handlungsplanung* nicht. Mit Blick auf die Itemsets zur Deduktion sowie Korrelation wird der Einfluss des Semesters etwas deutlicher (Anhang, Tabelle 71).

Werden die Unterschiede zwischen den Semestern näher betrachtet, so zeigt sich, dass beim *korrelativen Schließen* signifikante Unterschiede zwischen dem ersten und sechsten Semester vorliegen. Sechstsemester lösen die Items besser, es zeigen sich mittlere Effektstärken (Mittlere Differenz=-.43, SD=.13, $p=.004$; $d=.4$, $p<.01$). Nicht bedeutsam sind die Unterschiede zwischen dem ersten und dritten Semester (Mittlere Differenz=-.16, SD=.15, $p=.86$) sowie dem dritten und sechsten Semester (Mittlere Differenz=-.27, SD=.15, $p=.22$) (Mittelwertunterschiede Anhang, Tabelle 66).

Werden die kognitiven Anforderungen beim induktiven und korrelativen Schließen verglichen, so zeigt sich, dass beide ähnliche Fähigkeiten erfordern. Allerdings steht lediglich die Fähigkeit zum *korrelativen Schließen*, nicht aber die Fähigkeit zur Induktion, im Rahmen der Studie im Zusammenhang mit Expertise. Es stellt sich dementsprechend die Frage, inwieweit sich die Testaufgaben unterscheiden.

In formaler Hinsicht ist feststellbar, dass die Testaufgabe zur Induktion etwas leichter ist als das Variablenset zur Korrelation ($P_{\text{Induktion}}=58$; $P_{\text{Korrelation}}=51$). In Bezug auf die kognitive Anforderung erfordern beide die Kompetenz Einschränkungen zu erkennen. Sie unterscheiden sich darin, dass induktives Schließen die Vermeidung ungerechtfertigter Übergeneralisierungen, korrelatives Schließen die Vermeidung illegitimer Verknüpfungen zweier Variablen erfordert. Sowohl induktives als auch korrelatives Schließen basiert zwar auf Wenn-dann bzw. Je-Desto-Beziehungen. Korrelatives Schließen erfordert jedoch zudem die Berücksichtigung der Häufigkeit bzw. der Intensität der Wenn-dann bzw. Je-Desto-Beziehung. Zusätzlich zur Vermeidung illegitimer Schlüsse kommt dementsprechend hinzu, dass die Stärke der Beziehung zweier Variablen mitberücksichtigt werden muss. Es scheint, dass diese Anforderung mit zunehmender Studierenerfahrung besser bewältigt werden kann.

Werden die beiden Items zum *korrelativen Schließen* verglichen, so zeigt sich ein weiterer Unterschied: Bei der illegitimen Verknüpfung im Rahmen der Testaufgabe „Neuwagen“ zeigen sich keine bedeutsamen Unterschiede, wobei die Odds Ratio zeigt, dass das Item für Sechstsemester leichter ist als für Erstsemester (Anhang, Tabelle 73). Die Vermeidung der kausalen Verknüpfung im Rahmen der Testaufgabe „Alkohol“ ist hingegen etwas schwieriger zu beantworten ($P_{\text{Alkohol}}=60$; $P_{\text{Neuwagen}}=77$), auch für Sechstsemester. Es zeigen sich signifikante Unterschiede zwischen den Semestern (Anhang, Tabelle 72). Die Odds Ratio zeigt, dass Sechstsemester gegenüber Dritt- und Erstsemestern mehr als doppelt so häufig die Testaufgabe richtig lösen (Anhang, Tabelle 73). Es leitet sich daraus die Frage ab, inwieweit die Kontextualisierung Einfluss nehmen könnte. Wird bei beiden Testitems deshalb die Zusammenhangelogik der zu verknüpfenden Variablen mitberücksichtigt, so zeigen sich Unterschiede in der Plausibilität des Zusammenhangs. Es ist zu vermuten, dass der Zusammenhang zwischen Neuwagenkauf und Verkehrsunfällen den Studierenden weniger plausibler erschien als der Zusammenhang zwischen Alkoholpreisen und Alkoholkonsum.

Bei den Testitems zum *deduktiven Schließen* zeigt sich im Hinblick auf die Unterschiede zwischen den Semestern ein umgekehrtes Bild: Erstsemester lösen die Items signifikant besser als Dritt- und Sechstsemester. Dabei unterscheiden sich Erstsemester lediglich signifikant von Sechstsemestern (Mittlere Differenz=.39 SD=.14, $p=.011$). Zwar lösen sie auch gegenüber den Drittsemestern die Items besser, jedoch nicht signifikant (Mittlere Differenz=.30, SD=.15, $p=.14$).

Auf Itemsetebene zeigen sich bedeutsame Unterschiede beim schwierigsten Item „Gesundheit“ sowie beim leichtesten Testitem „Familienunternehmen“. Auch bei der Lösung des Testitems

„Berufstätigkeit“ sind Erstsemester tendenziell besser, allerdings nicht signifikant (Anhang, Tabelle 72, Tabelle 74).

... und die Rolle epistemologischer Überzeugungen

Die mehrfaktorielle Varianzanalyse erlaubt die Analyse von Haupt- und Interaktionseffekten. Es zeigt sich, dass epistemologische Überzeugungen (Sicherheit, Quelle) und Semester auf induktives ($R^2_{\text{kor}}=.02$, $F=2,65$, $df=4$, $p=.04$), deduktives ($R^2_{\text{kor}}=.05$, $F=4,72$, $df=4$, $p=.001$) sowie korrelatives Schließen ($R^2_{\text{kor}}=.09$, $F=7,38$, $df=4$, $p=.000$) signifikant Einfluss nehmen.

Das Modell klärt jedoch mit $R^2_{\text{kor}}=.09$ lediglich für korrelatives Schließen einen bedeutsameren Teil der Varianz. Wird deshalb korrelatives Schließen herausgegriffen, so zeigt sich, dass ausschließlich der Faktor der subjektiven Überzeugung der Sicherheit des Wissens einen signifikanten Beitrag leistet (Pillais Spur: $F=14,227$; $df=1$; $p=.000$). Der Einfluss des Semesters wird bei Berücksichtigung der epistemologischen Überzeugung der Sicherheit des Wissens irrelevant (Sicherheit: $F=3.183$; $df=6/252$; $p=.005$; Semester: $F=2,215$; $df=1$; $p=.111$). Eine reflektierte Überzeugung in Bezug darauf, dass Wissen auch unabhängig von Experten und Autoritäten weitergegeben werden kann, nimmt keinen bedeutsamen Einfluss (Quelle: $F=.045$, $df=1$, $p=.816$).

Zusammenfassend kann festgehalten werden, dass die Studienerfahrung bei einer quantitativen Betrachtung kaum Einfluss auf fächerübergreifende Problemlösekompetenzen zu haben scheint. Bemerkenswert ist etwa, dass sich Sechstsemester in ihrer Fähigkeit zum induktiven Schließen, Unterscheiden von Fakten und Meinung, der Zielevaluation und der Handlungsplanung nicht von Erst- und Drittsemestern unterscheiden, beim deduktiven Schließen waren Erstsemester den Sechst- und Drittsemestern sogar überlegen.

Bei einer Inbezugsetzung der kognitiven Teilfähigkeiten des induktiven und korrelativen Schließens zeigt sich, dass die Fähigkeit zum induktiven Schließen ein Prädiktor für die Leistung beim korrelativen Schließen ist ($R^2=.17$; $F=57,12$; $df=1$; $\beta=.41$; $T=7,56$; $p=.000$). Lediglich beim korrelativen Schließen, das heißt, bei einer Berücksichtigung der Intensität, bzw. Häufigkeit der gegebenen Wenn-Dann Beziehung, wird die Expertise im Sinne semesterbezogener Studienerfahrung relevant. Bei einer qualitativ-inhaltlichen Analyse ließ sich weiterhin die Relevanz der Plausibilität der zu verknüpfenden Variablen feststellen. Bedeutend bei der Frage des Einflusses des Semesters im Zusammenhang mit korrelativem Schließen ist die Berücksichtigung der subjektiven Überzeugungen der Studierenden. Es konnte gezeigt werden, dass bei Berücksichtigung der metakognitiven Komponente der Einfluss der Studienerfahrung obsolet zu werden scheint.

Mit Blick auf die Hypothesen ist festzuhalten, dass Hypothese 9, „Expertise [nimmt] Einfluss [...] und [...] Studierende eines höheren Semesters und Studierende eines niedrigeren Semesters [unterscheiden sich] im Hinblick auf ihre Problemlösekompetenz [...]“. Es wird angenommen, dass Studierende eines höheren Semesters eine höhere Problemlösekompetenz haben als Studierende eines niedrigeren

Semesters“, deshalb nur eingeschränkt angenommen werden kann. Hypothese 10, „diese Unterschiede [werden] vor allem bei schwierigen Aufgaben deutlich [...]“, muss differenziert werden und kann lediglich vor dem Hintergrund der Definition der Schwierigkeit als Logik des Zusammenhangs zweier Variablen angenommen werden.

Der Einfluss der Studiendisziplin

T-Tests zeigen, dass beim *induktiven* ($d=.2$, $p<.05$) und *korrelativen* ($d=.2$, $p<.05$) *Schließen* sowie im Subtest *Handlungsplanung* ($d=.3$, $p<.05$) signifikante Unterschiede zwischen Studierenden der Geistes- und der Sozialwissenschaften bestehen, die Disziplin nimmt mit kleinen Effekten Einfluss (Anhang, Tabelle 75). Der Blick auf die Mittelwerte zeigt, dass die Studierenden der Sozialwissenschaften die entsprechenden Itemsets besser lösen (Anhang, Tabelle 76).

Wird auch hier auf Ebene der Items differenziert, so wird deutlich, dass die signifikanten Unterschiede bei den Variablensets zum *induktiven* und *korrelativen Schließen* erst auf Variablensetebene durch die Aggregation der Leistung bei den Einzelitems zustande kommen. Auf Ebene der Einzelitems lösen zwar Sozialwissenschaftsstudierende die Items zum korrelativen und induktiven Schließen fast doppelt so häufig. Allerdings ist mit $\phi=.10-.11$ der Einfluss der Disziplin bei den Items „Alkohol“, „Migration“ und „Handwerker“ nur auf einem 90%-Niveau signifikant (Anhang, Tabelle 77). Konstruktionsbezogen wurden die Unterschiede zwischen den Items zum korrelativen Schließen bereits dargestellt. Die beiden Items zur Induktion unterscheiden sich insofern, dass mit dem Migrationsitem ein den Studierenden bekannter Kontext (Übergeneralisierung einer einzelnen Abiturnote), mit den Handwerkern ein den Studierenden weniger bekannter Kontext (Übergeneralisierung einer Einzelleistung im Handwerk) vorgegeben wurde. Befunde zum induktiven Schließen zeigen, dass die Erfahrung in bedeutsamem Maß Einfluss auf die Vermeidung illegitimer Übergeneralisierungen nimmt (Kap. 5). Mit Blick auf die vorliegenden Befunde ist deshalb weiter zu prüfen, inwieweit die Bekanntheit, bzw. Nähe des Kontexts hilfreich bei der Vermeidung der illegitimen Übergeneralisierung ist.

Bei einer Analyse auf Itemebene lässt sich weiter feststellen, dass Unterschiede zwischen den Disziplinen im Unterscheiden von subjektiver und objektiver Meinung zu erkennen sind. So lösen Sozialwissenschaftsstudierende das Item „Arbeitslose“ ($\phi=.13$; $p<.05$) signifikant häufiger, das Item „Raucher“ war für beide Studiengruppen schwieriger zu bearbeiten (die Sozialwissenschaftsstudierenden lösten das Item in etwa gleich schlecht wie die Geisteswissenschaftsstudierenden, Anhang, Tabelle 77). Die Frage ist auch hier, inwieweit sich die beiden Items unterscheiden: Beim Arbeitslosen-Item ist die subjektive Informationsquelle ein persönliches Statement einer Person, welches der objektiven Information – einer Statistik eines „offiziellen“ Institutes – widerspricht. Dabei ist die subjektive Information im Einklang mit dem zu bewertenden Statement formuliert. Die Statistik steht im Widerspruch zu dem zu bewertenden

Statement, wobei diese Information sehr eindeutig und leicht zu erkennen konzipiert ist (sinkende Arbeitslosenzahlen in der Statistik gegenüber steigenden Arbeitslosenzahlen in Statement und subjektiver Informationsquelle). Im Gegensatz hierzu ist das Item zum Sinken der Zahl der jugendlichen Raucher schwieriger konzipiert: Auch hier steht die persönliche Meinung in Einklang mit dem zu bewertenden Statement. Die Statistik erfordert jedoch eine genauere Analyse, bei einem flüchtigen Blick scheint sie zunächst der subjektiven Meinung zu entsprechen. Diese Einschränkung stellt eine zusätzliche Schwierigkeit dar.

Im Subtest Handlungsplanung gehen Studierende der Sozialwissenschaften mit einem geschlossenen, jedoch komplexerem Problem besser um als Geisteswissenschaftsstudierende. Das heißt, sie zerlegen Komplexität besser und setzen relevante Einzelschritte zur Problemlösung bedeutsam besser zusammen. Eine Fehleranalyse der Subtests Handlungsplanung und Zielevaluation erfolgt im nächsten Kapitel.

... und die Rolle epistemologischer Überzeugungen

Auch im Zusammenhang mit der Frage des Einflusses der Studiendisziplin wurden Haupteffekte untersucht. Dabei wurden lediglich die epistemologischen Überzeugungen und die Studiendisziplin in ihrem Zusammenhang mit Teilkomponenten fächerübergreifenden Problemlösens untersucht. Als abhängige Variablen wurden auch hier wieder die z-transformierten Testleistungen auf Ebene der Variablensets und der Subtests mit einbezogen.

Über den Einbezug der beiden subjektiven Überzeugungen als Kovariaten rückt auch hier der Einfluss der Disziplin in den Hintergrund und es lassen sich nur signifikante Einflüsse epistemologischer Überzeugungen bei induktivem ($F=3,36$, $df=3$, $p=.02$), deduktivem ($F=5,07$, $df=3$, $p=.00$) und korrelativem Schließen ($F=8,67$, $df=3$, $p=.00$) und der Handlungsplanung ($F=3,51$, $df=3$, $p=.02$) feststellen. Insgesamt klärt das Modell in überzufälligem Maß nur einen kleinen Teil der Testleistungsvarianz (Induktion: $R^2=.03$; Deduktion: $R^2=.05$; Korrelation: $R^2=.08$; Handlungsplanung: $R^2=.03$). In höherem Maß nimmt der Faktor der Sicherheit des Wissens lediglich Einfluss auf korrelatives Schließen ($R^2=.08$, $F=14,44$, $df=1$, $p=.00$; Pillais Spur: $F=3,08$, $df=6/253$, $p=.006$). Autoritätshörigkeit ($F=.75$, $df=6/253$, $p=.61$) und Disziplin ($F=95,df=6/253$, $p=.46$) nehmen keinen bedeutsamen Einfluss.

Abschließend lässt sich damit feststellen, dass ohne Berücksichtigung der epistemologischen Überzeugungen die Studierenden der Sozialwissenschaften bei der Vermeidung übergeneralisierender und kausaler Schlüsse sowie bei der Zerlegung komplexer Problemstellungen bedeutsam bessere Leistungen zeigen. Expertise – verstanden als Studiendisziplin – scheint Einfluss zu nehmen. Neben den Sozialisationseffekten ist jedoch auch hier der Einfluss der epistemologischen Überzeugungen zu berücksichtigen. Überwiegend zeigen sich signifikante Unterschiede in Abhängigkeit der

Einschätzung, für wie sicher, bzw. flexibel, Studierende Wissen halten. Das Zusammenspiel der epistemologischen Überzeugungen wird insbesondere im Hinblick auf die Handlungsplanung relevant. Der Einfluss der epistemologischen Überzeugungen steht auch hier wieder im Vordergrund.

Hypothese 11, „Unterschiede zwischen den Disziplinen [sind] beobachtbar [...]. Es wird angenommen, dass Sozialwissenschaftsstudierende, bei denen eher davon ausgegangen wird, dass sie im Rahmen ihres Studiums überwiegend mit Fragestellungen konfrontiert sind, im Rahmen derer illegitime Kausalitäten vermieden werden müssen, diejenigen Problemstellungen besser lösen können, die Kompetenzen des induktiven und korrelativen Schließens erfordern“, kann somit angenommen werden.

15.3 Fehleranalyse mit Blick auf Zielevaluation und Handlungsplanung

Im Folgenden wird die Leistung der Studierenden in den beiden Subtests Zielevaluation und Handlungsplanung beschrieben. Muster fehlender Werte bei beiden Subtests werden herausgegriffen und „Folgefehler“ berichtet. Insgesamt werden Fehlerstrukturen auch unter Berücksichtigung der Leistung beim logischen Schließen im fächerübergreifenden Kontext beschrieben.

Muster bei fehlenden Werten

Werden fehlende Werte betrachtet, so zeigt sich, dass die Nicht-Lösung eines der beiden Testitems überzufällig häufig mit der Nicht-Lösung des zweiten Items zusammenhängt. Diejenigen, die im Subtest Zielevaluation das Testitem „Verwaltungsapparat“ nicht bearbeiteten, bearbeiten auch nicht das Testitem „Lohnnebenkosten“ ($\chi^2=141,424$, $df=16$, $\phi=.703$, $p=.000$). Gleiches zeigt sich auch bei einem Vergleich der Testitems zur *Handlungsplanung*. Diejenigen Studierenden, die das Testitem „Arbeitslosenzahlen“ nicht bearbeiteten, bearbeiteten auch nicht das schwierigere Testitem „Getränkehersteller“ ($\phi=.493$, $\chi^2=68,908$, $df=1$, $p=.000$). Über die Subtests hinweg zeigen sich keine bedeutsamen Zusammenhänge.

Schwierigkeiten bei Testitems zur Zielevaluation

Die Items zur Zielevaluation wurden so konstruiert, dass zwei Ziele zu bewerten sind: Beide Testitems müssen im Hinblick auf die Zielerreichung (zielführend) und mit Blick auf die Kostenneutralität (kostenneutral) evaluiert werden.

Die Kostenneutralität ist beim Testitem „Verwaltungsapparat“ darüber festzustellen, dass die im Gutachten gegebenen – durch die Verwaltungsumstrukturierung – eingesparten Kosten mit den anfallenden Folgekosten verglichen werden müssen. Die hohe Differenz der Summen setzt keine Rechenfähigkeiten voraus („Folgekosten durch die Umstrukturierung 350.000 Euro“; „Einsparungen

von 75.000 Euro“). Die Evaluation der Zielerreichung erfolgt durch die Verknüpfung komplexerer Informationen. So muss durch Vergleich der Informationen „Verwaltungsorganigramm“, „Richtlinien der Stadt“ und „Verwaltungsevaluation“ erkannt werden, dass das einzige Problem, das in der Verwaltungsevaluierung durch die Bürger genannt wird, durch die Umstrukturierung der Verwaltung verschärft würde (noch weniger Mitarbeiter).

Insgesamt basiert die Evaluierung der Kostenneutralität auf dem Erkennen der relevanten Information, ihrer korrekten In-Bezug-Setzung sowie dem Abgleich mit dem Kriterium der Kostenneutralität. Die Evaluation des Aspektes der Zielerreichung erfordert die Verknüpfung des Befragungsergebnisses, der Richtlinien der Stadt und des Verwaltungsorganigramms. Einschränkende Bedingungen müssen erkannt werden.

Beim zweiten Testitem „Lohnnebenkosten“ müssen ebenfalls die Kriterien Kostenneutralität und Zielerreichung evaluiert werden. Auch hier ist die Frage der Kostenneutralität einfacher konstruiert als die der Zielerreichung. Sie ist mit einem Blick auf die Statistik der Entwicklung der städtischen Investitionen zu erfassen. Der Unterschied gegenüber dem Testitem „Verwaltungsapparat“ ist, dass die Information einer Statistik entnommen werden muss und nicht nur in einem Text genannt wird. Zur Einordnung der Zielerreichung müssen die in der Statistik abgebildeten Entwicklungsverläufe nachvollzogen und mit Informationen, die in Textform gegeben sind, verknüpft werden. Die beiden Testitems unterscheiden sich weniger in der Quantität der zu verknüpfenden Variablen, sondern vielmehr in der Modalität der Information (Verwaltungsapparat: Text, Grafik; Lohnnebenkosten: Statistik). Die Befunde sollen dementsprechend Annahmen über den Einfluss unterschiedlichen Informationsmaterials erlauben.

Beide Items haben einen mittleren Schwierigkeitsgrad („Verwaltungsapparat“ $P_i=46,5$; „Lohnnebenkosten“ $P_i=43,4$), ihre Lösung steht in Zusammenhang. Es ist zu beobachten, dass 70% derjenigen, die das Testitem „Verwaltungsapparat“ lösten, auch das Testitem „Lohnnebenkosten“ lösen konnten, bzw. umgekehrt ($\phi=.70$, $\chi^2=141,42$, $df=16$, $p=.000$).

Weiterhin zeigt sich, dass die Darbietungsmodalität eine Rolle zu spielen scheint: Die Evaluation des Faktors der Kostenneutralität beim Testitem „Verwaltungsapparat“ stellt kaum eine Schwierigkeit dar ($P_i=89,6\%$). Die Bewertung des Kriteriums Zielerreichung ist mit größeren Schwierigkeiten verbunden ($P_i=46,5$). Semester und Studiendisziplin unterscheiden sich dabei jedoch nicht bedeutsam. Beim Testitem „Lohnnebenkosten“ zeigt sich, dass die Evaluation des Kriteriums Kostenneutralität eine höhere Anforderung als beim Item „Verwaltungsapparat“ darstellt ($P_i=59,4$). Die Bewertung der Zielerreichung ist mit $P_i=62,9$ leichter. Auch hier unterscheiden sich Semester und Disziplinen nicht statistisch bedeutsam.

Werden die zu evaluierenden Kriterien zueinander in Beziehung gesetzt, so zeigt sich ein schwacher signifikanter Zusammenhang zwischen der korrekten Evaluation des Kriteriums Zielerreichung bei den Testitems „Verwaltungsapparat“ und „Lohnnebenkosten“ ($\phi=.17$; $\chi^2=7,06$; $df=1$, $p=.008$). Das

bedeutet, diejenigen Studierenden, die bei der Evaluation des Kriteriums der Zielerreichung des einen Items erfolgreich waren, konnten ebenfalls das Kriterium der Zielerreichung des zweiten Items richtig bewerten.

Wird die Fähigkeit zur Evaluation des Kriteriums der Zielerreichung mit der Fähigkeit zur Handlungsplanung in Beziehung gesetzt so zeigen sich signifikante Zusammenhänge für das schwierigere Item der „Lohnnebenkosten“ und Handlungsplanung, nicht mehr signifikant sind sie für das etwas leichtere Item der „Verwaltungsplanung“ und Handlungsplanung (Lohnnebenkosten_{ziel}-Handlungsplanung: $\chi^2=31,28$, $df=10$, $\eta=.35$, $p=.001$; Verwaltungsplanung_{ziel}-Handlungsplanung $\chi^2=16,54$, $df=10$, $\eta=.25$, $p=.08$). Die Evaluation des Kriteriums der Kostenneutralität steht nicht in überzufälligem Zusammenhang mit der Handlungsplanung (Lohnnebenkosten_{Kostenneutr}-Handlungsplanung: $\chi^2=8,14$, $df=10$, $\eta=.18$, $p=.62$; Verwaltungsplanung_{Kostenneutr}-Handlungsplanung $\chi^2=9,59$, $df=10$, $\eta=.19$, $p=.48$). Das bedeutet, die Subtests *Zielevaluation* und *Handlungsplanung* scheint insbesondere die Fähigkeit zu verbinden, unter Berücksichtigung einschränkender Bedingungen Informationen schlussfolgernd verknüpfen zu können.

Schwierigkeiten bei Testitems zur Handlungsplanung

Beide Testitems stellen die schwierigsten Testaufgaben dar. Zur Lösung müssen Maßnahmen in die richtige Reihenfolge gebracht werden. Jeweils sechs Informationsbausteine beschreiben die Ausgangssituation, zusätzlich enthält diese Information vier einschränkende Bedingungen. Den Studierenden liegt ein Maßnahmenkatalog mit fünf Maßnahmen (A-E) vor, die vor dem Hintergrund der Ausgangssituation so platziert werden müssen, dass das genannte Ziel erreicht wird. Aus diesen fünf Maßnahmen sind weiterhin zwei Maßnahmen für die Zielerreichung irrelevant. Von diesen beiden Distraktoren ist ein Distraktor distal konzipiert, ein zweiter ebenso irrelevant, jedoch aufgrund der Begrifflichkeit zu den Maßnahmen passender, das heißt als proximaler Distraktor konzipiert (bspw. Lohnnebenkosten vs. Löhne senken). Beide Distraktoren müssen erkannt und gestrichen werden. Beim Testitem „Getränkehersteller“ müssen soziale und ökologische Informationen zueinander in Bezug gesetzt werden (bspw. Naturschutz, Kindertagesstättenbau), beim Testitem „Arbeitslosigkeit“ wirtschaftsbezogene (bspw. Lohnnebenkosten, Unternehmensbilanzen). Die Schwierigkeitsindices zeigen, dass es geringere Anforderungen stellte, mindestens einen Distraktor zu erkennen, es hingegen schwieriger war, beide zu erkennen (Anhang, Tabelle 83). Der proximale Distraktor wurde häufiger als der distale Distraktor erkannt. Erklärt werden könnte dies darüber, dass der proximale Distraktor eher über die gegebene Information ausgeschlossen werden kann als der distale Distraktor, der als irrelevanterer Distraktor mit mehr Unsicherheiten verbunden ist.

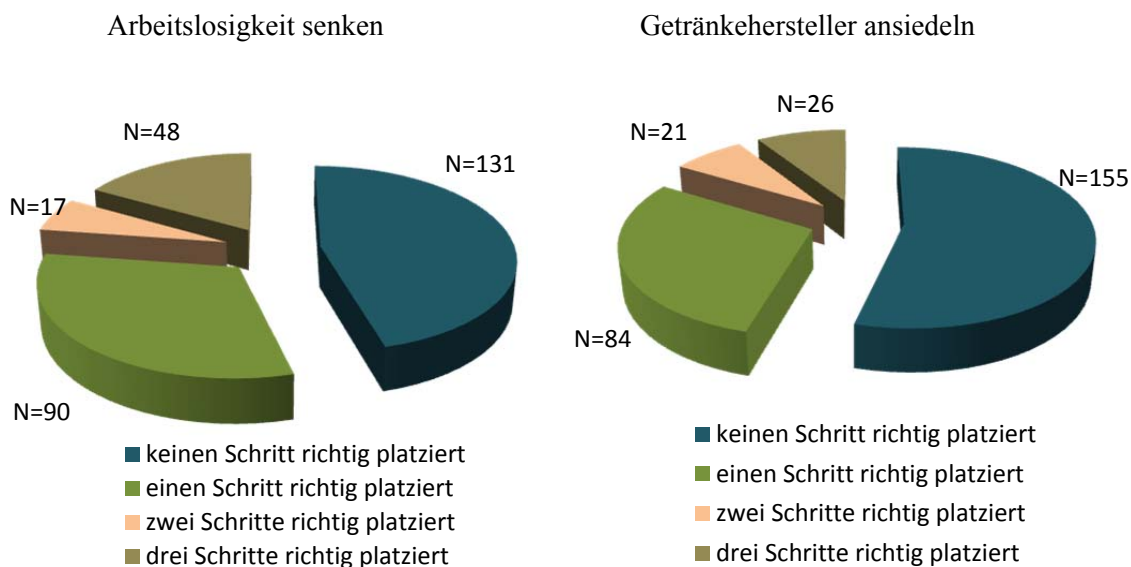
Die Analyse der korrekten Platzierung der Handlungsmaßnahmen zeigt, dass für beide Testaufgaben die Schwierigkeitsindices von $P_{iArbeitslosigkeit}=54,19$ bzw. $P_{iGetränkehersteller}=45,80$ im Mittel liegen. Dies gilt jedoch nur, wenn die korrekte Lösung definiert wird als „mindestens ein Schritt wurde korrekt

platziert“. Der Schwierigkeitsgrad erhöht sich drastisch, wenn lediglich diejenigen Studierenden betrachtet werden, die sowohl alle relevanten Maßnahmen richtig gereiht sowie beide Distraktoren korrekt erkannt haben (vgl. Pilotierung).

Die Schwierigkeit der korrekten Platzierung

Bei einer Analyse der richtigen Platzierung zeigt sich, dass insgesamt N=48 (Arbeitslosigkeit) bzw. N=26 (Getränkehersteller) Studierende alle drei Maßnahmen, also alle drei Handlungsschritte, korrekt platzierten (Grafik 15, Anhang, Tabelle 78). Von diesen erkannten N=10 (Arbeitslosigkeit), bzw. N=24 (Getränkehersteller) auch beide Distraktoren korrekt. Die Gesamtlösung der beiden Testitems korreliert mit $r=.31$ ($p<.01$). Das Erkennen der Distraktoren korreliert mit der richtigen Platzierung im Falle des Items „Getränkehersteller“, weniger im Falle des Items „Arbeitslosigkeit“ ($r_{\text{Getränke}}=.25$, $p<.01$; $r_{\text{Arbeitslosigkeit}}=.14$, $p<.01$). Das bedeutet, die Anforderung, Distraktoren zu erkennen, lässt nicht unbedingt darauf schließen, dass auch die Maßnahmen korrekt platziert werden können. Das Erkennen der Distraktoren bei beiden Testitems hängt in mittlerem Maße signifikant zusammen, das heißt, diejenigen Studierenden, die die Distraktoren eines Items erkannten, erkannten in statistisch bedeutsamem Maß auch die Distraktoren des zweiten Items ($r=.34$, $p<.01$).

Grafik 15: Richtige Platzierung der Maßnahmen



Werden die häufigsten Lösungsschritte deskriptiv betrachtet, so lassen sich die Studierenden weiter nach der Anzahl der nicht berücksichtigten Einschränkungen unterscheiden (Anhang, Tabellen 79-82).

15.4 Zusammenhänge zwischen Teilkompetenzen fächerübergreifenden Problemlösens

Nach der Beschreibung fächerübergreifender Problemlösekompetenzen wird abschließend dargestellt, inwieweit Zusammenhänge zwischen kognitiven Teilkompetenzen festgestellt werden können. Zur Messung fächerübergreifender Problemlösekompetenzen wurden kognitive Teilkompetenzen operationalisiert, die im Rahmen des Problemlöseprozesses relevant werden. Bislang fokussierte die Darstellung der Befunde Teilkompetenzen. Um eine komplexe fächerübergreifende Problemstellung erfolgreich lösen zu können, werden jedoch nicht einzelne Teilfähigkeiten relevant, vielmehr ist ihr Zusammenwirken notwendig. Die Annahme der vorliegenden Studie ist deshalb, dass die operationalisierten Teilprozesse des Problemlösens als operative Teilkomponenten auch in ihrem Zusammenhang nachgezeichnet werden können. Das bedeutet, die Teilkompetenzen sollten mit Blick auf die ihnen jeweils zugrundeliegenden kognitiven Fertigkeiten sowie den Anforderungsgehalt der Testaufgaben in Zusammenhang gebracht werden können. Weiterhin wurde den Subtests eine zunehmende Komplexität unterstellt. Das bedeutet, die Subtests der Zielevaluation und Handlungsplanung erfordern kognitive Fähigkeiten, die im Rahmen des Subtests Situationsanalyse gemessen werden. Ebenso wird davon ausgegangen, dass die Kompetenzen zur Handlungsplanung kognitive Fähigkeiten der Zielevaluation erfordert und umgekehrt (Vergleich und Bewertung von Informationen, Zielverfolgung).

Die im Rahmen der Subtests konzipierten Testitems sollen folgende spezifische Teilkompetenzen erfassen:

Die Items zur Situationsanalyse erfordern Teilkompetenzen der

- Informationssammlung und -integration,
- des Erfassens der Zusammenhänge zwischen der gegebenen Information sowie
- der Hypothesenbildung zur abschließenden Beschreibung der jeweiligen Situation.
- Dabei werden die gegebenen Informationen über deduktives, induktives, bzw. korrelatives Schließen bzw. den Abgleich zwischen Fakten und Meinung verknüpft.

Die Testitems zur Zielevaluation erfordern darüber hinaus die Teilkompetenzen

- des Vergleichs
- und der Bewertung der gegebenen Information
- sowie der Zielverfolgung.

Die Testaufgaben der Handlungsplanung erfordern weiterhin

- komplexe Zerlegungsprozesse,
- Erkennen multipler Einschränkungen und
- Entscheidung im Hinblick auf die Maßnahmenreihung.

Damit legt die theoretische Konzeption des Tests Zusammenhänge nahe, die im Folgenden überprüft werden.

Können entsprechende Beziehungen beschrieben werden? Folgende korrelative Verknüpfungen zeigen sich zwischen Teilkompetenzen der Situationsanalyse (es werden nur Zusammenhangsstärken $r \geq .2$ berichtet, Anhang, Tabelle 84): Der stärkste Zusammenhang zeigt sich zwischen den Anforderungen des korrelativen und induktiven Schließens ($r = .41$, $p < .001$). Weiterhin sind Zusammenhänge zu beobachten zwischen der Anforderung, korrelative Beziehungen zwischen Informationen zu erkennen und der Fähigkeit zwischen subjektiver und objektiver Information zu unterscheiden ($r = .17$, $p < .01$). Negativ korrelieren induktives und korrelatives Schließen mit deduktivem Schließen ($r_{\text{Induktion}} = -.16$, $p < .01$, $r_{\text{Korrelation}} = -.20$, $p < .01$). Die Fähigkeiten zur Zielevaluation sowie zur Handlungsplanung stehen in Zusammenhang mit der Fähigkeit zum Unterscheiden zwischen objektiver und subjektiver Information ($r_{\text{Handlungsplanung}} = .15$, $p < .01$; $r_{\text{Zielevaluation}} = .17$, $p < .01$), Handlungsplanung und Zielevaluation korrelieren mit $r = .24$ signifikant ($p < .01$).

Für diese Zusammenhänge lassen sich regressionsanalytisch folgende Vorhersageklärungen finden: Die Fähigkeit, korrelative Schlüsse von kausalen zu unterscheiden, wird überwiegend über die Fähigkeit zum induktiven Schließen erklärt ($R^2 = .19$, $T = 7,46$, $\beta = .40$, $p < .01$). Studierende, die korrelative von kausalen Zusammenhängen gut unterscheiden können, zeichnen sich weiterhin auch durch die Fähigkeit aus, komplexe Probleme (Handlungsplanung) besser zerlegen zu können und Fakten und Meinung unterscheiden zu können ($T_{\text{Handlungsplanung}} = 2,13$, $\beta = .12$, $p < .05$; $T_{\text{Faktens.Meinung}} = 2,12$, $\beta = .12$, $p < .05$). Die Fähigkeit, induktive Schlüsse korrekt zu erkennen, kann lediglich über die Fähigkeit zum korrelativen Schließen vorhergesagt werden ($R^2 = .17$, $T = 7,56$, $\beta = .41$, $p < .01$). Wird die Fähigkeit betrachtet, subjektive Information von objektiver unterscheiden zu können, so zeigt sich, dass 5% der Leistungsvarianz über die Fähigkeiten erklärt werden kann, illegitime kausale Schlüsse erkennen zu können sowie komplexe Probleme zerlegen und Handlungen planen zu können ($T_{\text{Handlungsplanung}} = 2,22$, $\beta = .13$, $p < .05$; $T_{\text{Korrelationvs.Kausalität}} = 2,56$, $\beta = .15$, $p < .05$). Die Fähigkeit zur Deduktion lässt sich für die Stichprobe nicht erklären. Teilkompetenzen der Situationsanalyse sowie Zielevaluation stehen mit Fähigkeiten zur Handlungsplanung insofern in Zusammenhang, dass die Vermeidung kausaler Schlüsse sowie Fähigkeiten der Zielevaluation und Zielbewertung die Fähigkeit erklären, komplexere Handlungen planen zu können ($T_{\text{Zielevaluation}} = 4,24$, $\beta = .24$, $p < .01$; $T_{\text{Korrelationvs.Kausalität}} = 2,35$, $\beta = .13$, $p < .05$). Die Fähigkeit, Ziele bewerten und evaluieren zu können lässt sich für die Stichprobe wiederum über die Fähigkeit zur Handlungsplanung erklären ($T = 4,18$, $\beta = .24$, $p < .01$).

Wie können die Anforderungen, die die Testitems stellen, in einen inhaltlichen Zusammenhang gebracht werden? Es zeigt sich, dass diejenigen Teilkompetenzen zusammenhängen, die sich theoretisch über die Fähigkeit beschreiben lassen, komplexere Einschränkungen zu erkennen. Die

Befunde zeigen, dass Studierende, die illegitime Übergeneralisierungen erkennen, weniger Schwierigkeiten haben, auch illegitime kausale Schlüsse zu erkennen. Dabei lässt sich empirisch auch zeigen, dass die Fähigkeit zum korrelativen Schließen komplexer zu sein scheint, als die Fähigkeit, induktive Schlüsse zu vermeiden. Die Fähigkeiten, komplexe Probleme zerlegen und Handlungsschritte planen zu können sowie Fakten und Meinung unterscheiden zu können, klären neben dem induktiven Schließen einen Teil der Leistungsvarianz beim korrelativen Schließen. Im Hinblick auf die Frage, inwieweit kognitive Fähigkeiten der Handlungsplanung und der Zielevaluation mit Fähigkeiten der Situationsanalyse in Zusammenhang gebracht werden können, zeigt sich, dass das verbindende Element das Erkennen von Einschränkungen bzw. der Umgang mit komplexeren Unsicherheiten ist. Im Zusammenhang mit der Anforderung der Handlungsplanung wird die Kompetenz zur Zielevaluation sowie die Fähigkeit zum korrelativen Schließen und zur Unterscheidung von Fakten und Meinung relevant. Die Fähigkeit zur Zielevaluation ist weniger über Fähigkeiten der Situationsanalyse zu erklären, sie lässt sich eher über Fähigkeiten der Handlungsplanung beschreiben.

Diese Zusammenhänge wurden im Rahmen verschiedener Strukturmodelle überprüft. Es zeigt sich, dass die beschriebene Kombinatorik kognitiver Kompetenzen beim fächerübergreifenden Problemlösen mit einem guten Modelfit bestätigt werden kann (Anhang, Tabellen 86-88). Die Zusammenhänge zwischen Kompetenzen der Situationsanalyse und Kompetenzen der Zielevaluation und Handlungsplanung sind eher schwach und lediglich über Zusammenhänge zwischen der Handlungsplanung und korrelativem Schließens, bzw. dem Unterscheiden von Fakten und Meinung zu modellieren. Dies bestätigt insgesamt auch ein Vergleich derjenigen Studierenden nochmals, die kognitive Anforderungen der Situationsanalyse im fächerübergreifenden Problemlösen besser bewältigen: Sie lösen lediglich die Testitems zur Handlungsplanung signifikant besser (Anhang, Tabelle 85). Dieser Zusammenhang lässt die Vermutung zu, dass die Komplexität planerischen und bewertenden Schließens im Rahmen der Zielevaluation und Handlungsplanung eher wenig mit Kompetenzen der Situationsanalyse insgesamt gemeinsam hat. Vielmehr müssen Teilkompetenzen berücksichtigt werden, da eher die Anforderung zum Erkennen einschränkender Bedingungen das verbindende Element darstellt. Die Zusammenhänge zwischen Zielevaluation und Handlungsplanung sind beidseitig bedeutsam (Modelle 2 und 3), wobei die Fähigkeit zur Handlungsplanung, wie auch die Regressionsanalyse bereits zeigte, mehr zur Erklärung der Testleistung bei der Zielevaluation beiträgt.

Im Hinblick auf die Hypothesen lässt sich demzufolge folgendes festhalten:

Hypothese 12, „kognitive Teilkomponenten fächerübergreifenden Problemlösens, die über ähnliche kognitive Anforderungen beschrieben werden können, [stehen] auch empirisch in Zusammenhang [...]“, kann für induktives und korrelatives Schließen sowie die Anforderung des Erkennens einschränkender Bedingungen (korrelatives Schließen, Fakten und Meinung unterscheiden, Anforderungen der Handlungsplanung) angenommen werden.

Hypothese 13, „kognitive Teilkompetenzen zur Analyse von Problemsituationen (Subtest Situationsanalyse) [stehen] in Zusammenhang [...] mit Kompetenzen der Zielevaluation und Handlungsplanung“, kann mit Blick auf Zusammenhänge zwischen korrelativem Schließen und dem Unterscheiden von Fakten und Meinung sowie dem Subtest Handlungsplanung angenommen werden. Hypothese 14, „Studierende, die Testitems des Subtests Situationsanalyse besser lösen, lösen auch die Testitems der Subtests Zielevaluation und Handlungsplanung besser“, kann nur eingeschränkt mit Blick auf die Zusammenhänge zwischen Teilkompetenzen der Situationsanalyse und der Handlungsplanung angenommen werden.

15.5 Gesamtmodell zur Vorhersage fächerübergreifender Problemlösekompetenz

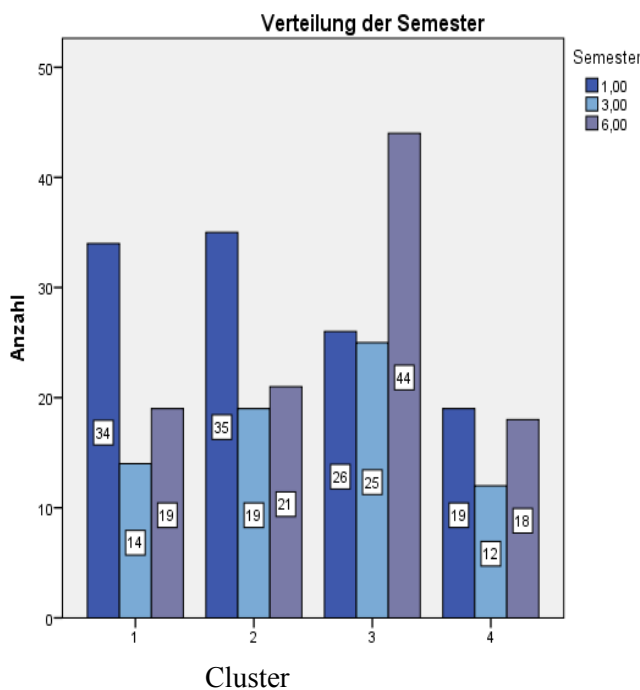
Nach der Beschreibung der Kompetenzstrukturen und der Erarbeitung von Vorhersagemodellen sowie der Kombination kognitiver Teilkompetenzen werden die Befunde abschließend in einem Gesamtmodell zusammengefasst. Die Differenzierung über homogene Gruppen verdeutlicht, inwieweit aus der heterogenen Gesamtheit fächerübergreifender Teilkompetenzen und motivationaler, affektiver, metakognitiver und expertisebezogener Kovariaten Teilmengen identifiziert werden können, die sich statistisch bedeutsam unterscheiden (Backhaus, Erichson, Wulff & Weiber, 2000).

Die Clusterzentrenanalyse für die z-transformierte Testleistung zeigt, dass insgesamt 4 Gruppen bedeutsam unterschieden werden können (Anhang, Tabellen 89-93): Die homogenen Studierendengruppen sind etwa gleich groß mit Cluster 1 N=67 Studierenden, Cluster 2 N=75, Cluster 3 N= 95 und Cluster 4 N=49 Studierenden. Die Gruppen unterscheiden sich signifikant im Hinblick auf korrelatives und induktives Schließen sowie die Fähigkeit zur Handlungsplanung (Induktion: $\chi^2=164,59$, $df=6$, $p<.01$, $\eta=.73$; Korrelation: $\chi^2=40,20$, $df=6$, $p<.01$, $\eta=.65$; Handlungsplanung: $\chi^2=93,37$, $df=30$, $p<.01$, $\eta=.44$). Weiterhin bestehen signifikante Unterschiede mit Blick die Unterscheidung zwischen Fakten und Meinung und die Zielevaluation, ebenso mit Blick auf deduktives Schließen (Fakten und Meinung: $\chi^2=60,99$, $df=6$, $p<.01$, $\eta=.46$; Zielevaluation: $\chi^2=188,70$, $df=69$, $p<.01$, $\eta=.64$; Deduktion: $\chi^2=143,39$, $df=9$, $p<.01$, $\eta=.65$).

Ein Vergleich zeigt weiter, dass sich die Studierendengruppen in Abhängigkeit der Ausprägung metakognitiver und motivationaler Kovariaten in ihren Teilkompetenzen des fächerübergreifenden Problemlösens unterscheiden (Grafik 18, S.156; Anhang, Tabellen 91-92). Cluster 3 fasst die reflektiertesten Studierenden mit der ausgeprägtesten Überzeugung im Hinblick auf die Wirksamkeit des eigenen Handelns zusammen, Cluster 2 diejenigen Studierenden, die ebenso reflektiert sind, jedoch im Vergleich zu Cluster 3 signifikant weniger überzeugt von der Wirksamkeit ihres Handelns. Die beiden Studierendengruppen unterscheiden sich dementsprechend nicht im korrelativen und induktiven Schließen, sondern in den Teilkompetenzen der Handlungsplanung, Zielevaluation und der Unterscheidung von Fakten und Meinung. Die Cluster 1 und 4 fassen die im Hinblick auf Wissen und

Wissensentwicklung am wenigsten reflektierten Studierenden zusammen. Die Studierenden des vierten Clusters unterscheiden sich von den reflektiertesten Studierenden des dritten Clusters in Bezug auf beide gemessenen epistemologischen Überzeugungen, die Studierenden des ersten Clusters lediglich mit Blick auf die Einschätzung der Sicherheit des Wissens. Die Studierenden des vierten Clusters zeigen bei allen Testitems signifikant schlechtere Leistungen, die Studierenden des Clusters 1 ausschließlich beim induktiven und korrelativen Schließen. Die Studiengruppen unterscheiden sich nicht mit Blick auf das am Studienfach gezeigt Interesse sowie die Einschätzung des Transfers fachspezifisch erlernter Problemlösemethoden auf fächerübergreifende Probleme.

Grafik 16: Semesterverteilung je Cluster

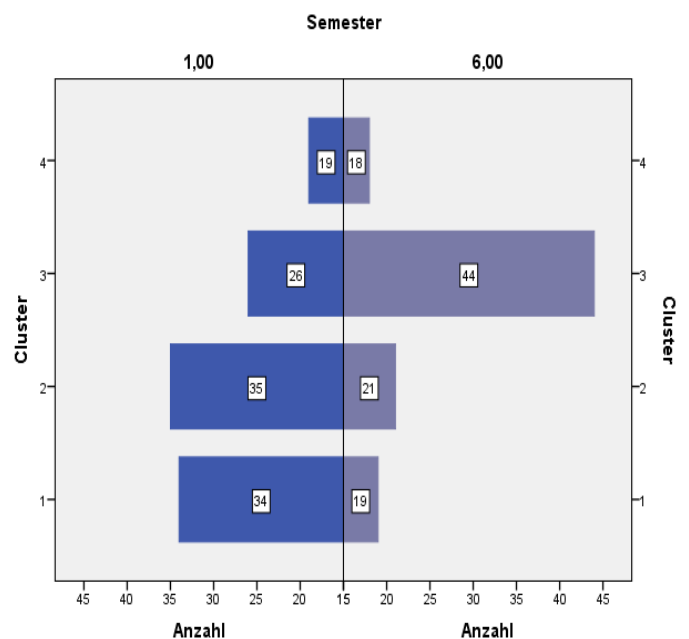


Die Studiengruppen unterscheiden sich jedoch signifikant im Hinblick auf das Studiensemester: die Studierenden des ersten und zweiten Clusters sind überwiegend im ersten, die Studierenden des dritten Clusters überwiegend im sechsten Semester (Grafik 16; $\chi^2=12,38$, $df=6$, $p<.05$, $V=.15$).

Grafik 17: Jüngere und ältere Semester je Cluster

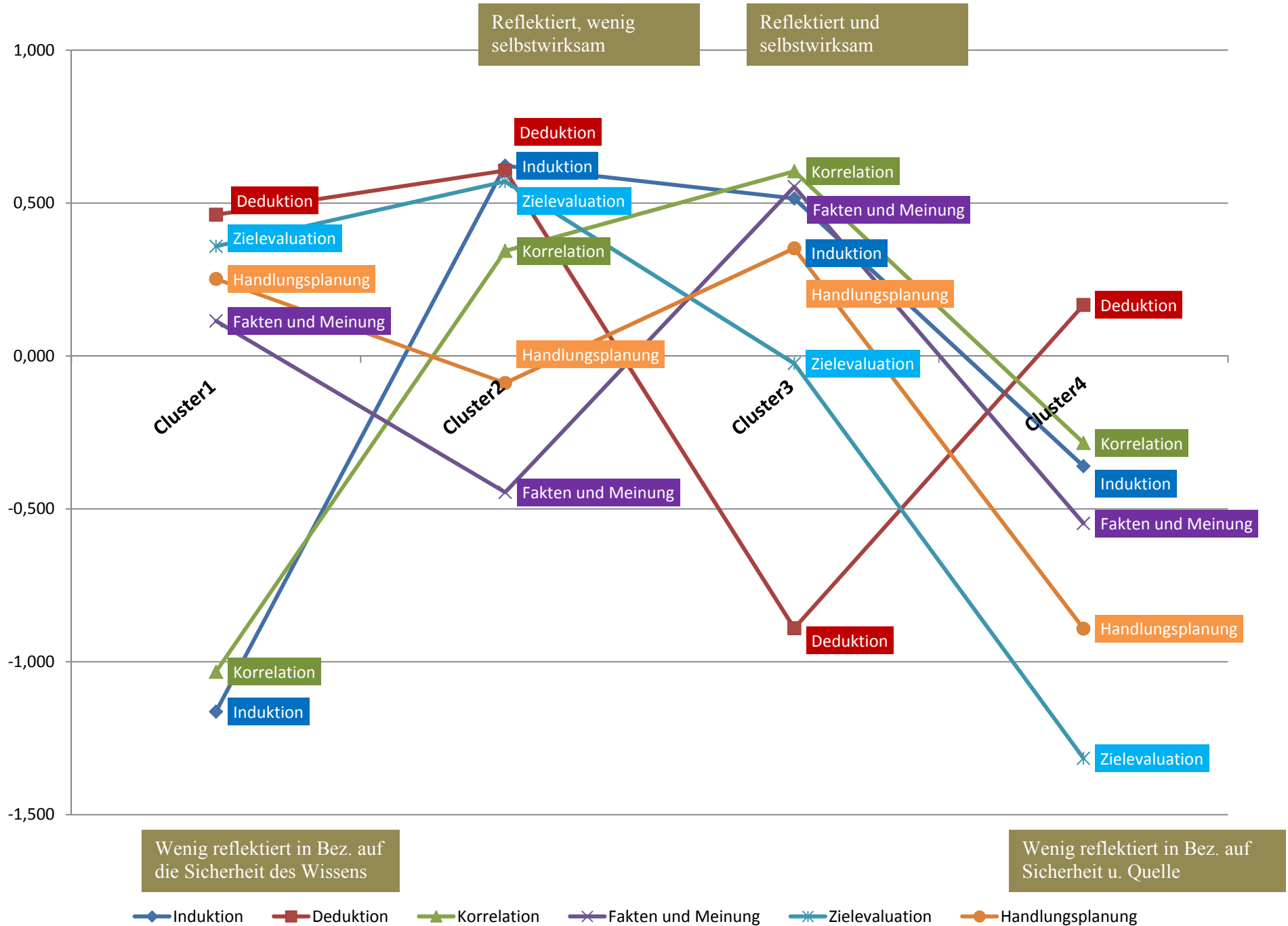
Ein Vergleich der Studiendisziplinen zeigt, dass im dritten Cluster mehr Sozialwissenschaftsstudierende zusammengefasst sind als Geisteswissenschaftsstudierende, jedoch nicht überzufällig häufig ($\chi^2=3,28$, $df=3$, $p>.05$, $V=.11$).

Damit lassen sich die Befunde der vorhergehenden Analysen folgendermaßen zusammenfassen: Reflektiertere Studierende zeigen insgesamt bedeutsam bessere Testleistungen insbesondere bei



denjenigen Testaufgaben, die auf der Einschätzung von Wahrscheinlichkeiten beruhen. Die Überzeugung von der Wirksamkeit des eigenen Handelns wird relevant bei Testitems, die den Umgang mit komplexen Unsicherheiten, weniger die Einschätzung von Wahrscheinlichkeiten erfordern. Insgesamt wenig reflektierte Studierende, das heißt solche Studierende, die bedeutsam weniger der Überzeugung sind, dass sich Wissen weiterentwickelt, bzw. dass Wissen auch über die eigene Auseinandersetzung oder die Auseinandersetzung in Studierendengruppen weiterentwickelt werden kann, lösen die Testitems insgesamt bedeutsam schlechter. Die Studienerfahrung, also die Expertise, die Studierende mit zunehmendem Semester erwerben, unterscheidet reflektierte und weniger reflektierte Studierendengruppen, allerdings lediglich im Hinblick auf die Überzeugung bezüglich der Sicherheit des Wissens, dabei muss jedoch auch die Überzeugung von der Wirksamkeit des eigenen Handelns mit berücksichtigt werden.

Grafik 18: Zusammenfassender Vergleich der Cluster



Teil V: Diskussion und Ausblick – Einordnung der zentralen Befunde

Das Ziel der Studie war, ein Testinstrument zum fachübergreifenden Problemlösen bei Studierenden zu entwickeln und zu evaluieren. Nach einer Analyse der Binnenstruktur des Konstrukts wurden die Kompetenzen der Studierenden über einen Vergleich des ersten, dritten und sechsten Semesters und der Studiendisziplinen der Sozial- und Geisteswissenschaften beschrieben. Schließlich wurden Zusammenhänge mit metakognitiven, motivationalen und affektiven Kovariaten modelliert, um differentialpsychologisch die Unterschiede genauer beschreiben zu können.

Der Entwicklung des Testverfahrens lagen kognitionspsychologische Annahmen und Befunde zum schlussfolgernden Denken und Problemlösen im fächerübergreifenden Kontext zugrunde. Herangezogen wurden heuristische Modelle des Problemlösens, um die für die zentralen Prozesse des Problemlösens relevanten kognitiven Fähigkeiten zu operationalisieren. Problemcharakteristika gaben Anhaltspunkte zur Konzeption der Schwierigkeit der Items. Es wurde definiert, was unter einem fächerübergreifenden Problem, bzw. fächerübergreifenden Kompetenzen zu verstehen ist. Weiterhin wurde dargestellt, welche Anforderungen an die Messung entsprechender Kompetenzen gestellt werden. Daneben dienten auch vorliegende testdiagnostische Instrumentarien, vor allem die PISA-Studien zum fächerübergreifenden Problemlösen, als Orientierung zur Konzeption des vorliegenden Instrumentes. Zur Identifikation von metakognitiven, affektiven und motivationalen Kovariaten fächerübergreifender Problemlösefähigkeit wurden diejenigen Faktoren herangezogen, die besonders bedeutsam bei der Erklärung von Leistungsunterschieden sind: Epistemologische Überzeugungen sowie Selbstwirksamkeit und Studieninteresse wurden demzufolge mit erhoben. Unter Einbezug dieser Faktoren wurden Bedingungen fächerübergreifenden Problemlösens und interindividuelle Unterschiede beschrieben.

Die Studie leistet insgesamt einen Beitrag zur Beschreibung kognitiver Kompetenzen beim Problemlösen im fächerübergreifenden Kontext und ermöglicht über den Einbezug der Kovariaten einen Blick auf ausgewählte Bedingungsvariablen. Das theoretische Modell wurde so konzipiert, dass die Ergebnisse der empirischen Überprüfung Ansatzpunkte für die Praxis der universitären Lehre erlauben. Auf diese wird im Rahmen der Diskussion eingegangen. Die empirische Überprüfung erfolgte anhand einer querschnittlich ausgerichteten Studie. Im Folgenden werden die zentralen Befunde diskutiert. Eine Kritik der Studie leitet auf weiterführende Überlegungen und einen Ausblick über.

16. Das Testverfahren zur fächerübergreifenden Problemlösekompetenz: Theoretisches Konzept, Güte des Testverfahrens und Binnenstruktur des Konstrukts

Problemlösungen verlangen einen Vergleich von Aussagen und Prämissen, um zu einem gültigen, das heißt logisch richtigen Schluss zu gelangen (u.a. Anderson, 2007; Buehner & Cheng, 2005; Evans, 2005; Mayer, 2002; Rips, 1994). Für das vorliegende Testverfahren wurden davon ausgehend zentrale

Formen logischen Schließens operationalisiert, die im Rahmen der Problemlösephasen der Problemanalyse, Handlungsplanung und Zielevaluation relevant werden. Den Testaufgaben liegen die kognitiven Anforderungen der Identifikation relevanter Variablen, ihres Vergleichs sowie des Erkennens von Beziehungen zugrunde. Es müssen Annahmen konstruiert und Einschränkungen berücksichtigt werden, um über die Abwägung zwischen Alternativen zu einer Lösung zu gelangen. Die Testaufgaben zur Situationsanalyse wurden als Testitems zum induktiven, deduktiven und korrelativen Schließen konstruiert und erheben kognitive Anforderungen, die im Rahmen der Analyse und Erfassung von Problemsituationen relevant werden. Testaufgaben zum logischen Schließen auf Basis des Unterscheidens zwischen Fakten und subjektiver Meinung ergänzen die Items zur Situationsanalyse. Testitems zur Zerlegung eines komplexeren Problems in Teilschritte im Rahmen des Subtests Handlungsplanung wurden konstruiert, um die kognitiven Anforderungen der Planung einer Problemlösung zu simulieren. Die Einordnung von Zielen auf Basis gegebener Informationen war Fokus des Subtests Zielevaluation. Im Zentrum dieser Items stand die Fähigkeit, Problemlösungen im Hinblick auf ihre Zielerreichung bewerten und einordnen zu können.

Werden die Befunde zur testtheoretischen Überprüfung des Testverfahrens rekapituliert, so zeigt sich, dass es gelang, Items unterschiedlicher Schwierigkeit zu konstruieren. Die Items mittlerer Schwierigkeit differenzieren gut zwischen besseren und schwächeren Problemlösern, die Items mit höherem und geringerem Schwierigkeitsniveau verdeutlichen Stärken und Schwächen der Studierenden. Die Testitems zur Differenzierung zwischen subjektiver und objektiver Information ebenso wie die Testitems zur Deduktion sind weniger konsistent und müssten dahingehend überarbeitet werden. Ebenso sind die Testitems zur Handlungsplanung auch nach der Pilotierung noch sehr schwierig. Allerdings erlauben die Testitems zur Handlungsplanung unterschiedliche Auswertungsmodi und damit differenzierte Informationen zu Kompetenzausprägungen der Studierenden. Die Reliabilitäten der Subtests, die für die Pilotierung berechnet wurden, zeigen, dass annehmbare bis gute erste Reliabilitäten zwischen $\alpha=.4$ (Zielevaluation) und $\alpha=.7$ (Situationsanalyse und Handlungsplanung) erzielt wurden. Die Kürzung des Instrumentes für die Hauptstudie sowie die Dichotomie der Items führten natürlich nicht zu einer weiteren Verbesserung der Reliabilität. Insgesamt erlaubte das konstruierte Testinstrument in dieser kurzen Form jedoch bereits, Kompetenzausprägungen in unterschiedlichen Situationen des fächerübergreifenden Problemlösens zu erheben und Bedingungen fächerübergreifender Problemlösekompetenz zu identifizieren. Weiterhin konnten in theoretischer Hinsicht Binnenstrukturen des Konstruktes identifiziert werden.

Die Befunde zur Binnenstruktur zeigen, dass die drei Subtests in Abhängigkeit der ihnen zugrundeliegenden Anforderungen gemeinsame Varianz teilen, allerdings kein gemeinsamer Faktor allgemeiner fächerübergreifender Problemlösekompetenz beschrieben werden kann. Der Befund bestätigt die der Arbeit zugrundeliegende Erwartung unterschiedlicher Teilkomponenten fächerübergreifenden Problemlösens und stimmt damit mit zahlreichen Studien zur Differenzierung

von Problemlösekompetenzen in Abhängigkeit der Komplexität des Problems überein (u.a. Baumert, Klieme, Neubrand et. al., 2000). Die Befunde zeigen, dass zwischen logischem Schließen im Rahmen der Situationsanalyse sowie kognitiven Fähigkeiten des evaluativen oder bewertenden Schließens als Grundlage für planerisches Handeln differenziert werden kann. Das Testverfahren erlaubt damit eine funktionale Beschreibung kognitiver Kompetenzen im Rahmen der Problemlösephasen der Situationsanalyse sowie der planerischen und bewertenden Problemlösung. Weiterhin können auch die Einzelitems entsprechend ihrer kognitiven Anforderung differenziert werden. Insgesamt kann das Testverfahren voneinander unterscheidbare Teilkompetenzen erfassen, die zur Bewältigung unterschiedlicher Phasen im Prozess des Problemlösens notwendig sind. Die Messung erlaubt damit eine mehrteilige Analyse und spiegelt die Mehrdimensionalität des Konstruktes im Hinblick auf die operationalisierten Teilprozesse des Problemlösens wider (Vonken, 2005; auch Klieme, Maag-Merki & Hartig, 2007; Klieme & Hartig, 2007; Klieme & Hartig 2006; Wilson, 2005).

Wie ist das Testverfahren dementsprechend in die bisherige Diagnostik fächerübergreifenden Problemlösens einzuordnen? Bei einem Vergleich mit der Messung fächerübergreifender Problemlösekompetenzen im Rahmen der PISA-Studien ist festzustellen, dass Teilkomponenten wie dynamisches Problemlösen oder die Erfassung und Anwendung spezifischer Regeln (bspw. „Raumfahrtspiel“) über das vorliegende Testverfahren nicht erfasst werden. Im Gegensatz zu den PISA-Studien lag der Fokus auf der differenzierten Erfassung von Kompetenzen, die insbesondere im Rahmen der Situationsanalyse sowie im Rahmen der Zielevaulation und Handlungsplanung relevant werden und den Umgang mit unterschiedlich schwierigen Informationen erfordern. In diesem Zusammenhang stellt das Testverfahren eine Ausdifferenzierung des Konzeptes fächerübergreifender Problemlösekompetenzen dar und erlaubt einen Blick auf die hierzu relevanten kognitiven Teilkompetenzen sowie auf die Abhängigkeit der Kompetenzen von der Variation der Schwierigkeit der gegebenen Information. Im Gegensatz zu Verfahren des kritischen Denkens, die sich an curricularen Inhalten orientieren oder Eignungstests, die sich eher am Konzept der Testintelligenz orientieren, erlaubt das vorliegende Instrument eine Erfassung derjenigen Teilkompetenzen des Problemlösens, die im Studium fachspezifisch vermittelt werden, im späteren beruflichen Alltag bei der Bewältigung neuer Probleme notwendig werden, jedoch dort in einem unbekanntem, neuartigen Kontext auftauchen.

17. Die Frage der Domänenspezifität fächerübergreifender Problemlösekompetenzen

Der Frage, inwieweit das Testverfahren fächerübergreifende Kompetenzen erfasst, wurde anhand der Analyse gemeinsamer Varianz der Testintelligenz und der Leistung in den Subtests sowie anhand der Frage beobachtbarer Unterschiede zwischen den Disziplinen nachgegangen.

Die Befunde zeigen, dass deduktives Schließen und Testintelligenz über gemeinsame Anforderungen beschrieben werden kann. Weiterhin zeigt sich, dass die Anforderungen des induktiven, bzw. korrelativen und bewertenden Schließens von deduktivem Schließen und Testintelligenz separiert werden können und dass auch planerisches und bewertendes Schließen nicht über einen gemeinsamen Faktor mit deduktivem Schließen sowie Testintelligenz beschrieben werden kann. Weiter konnte gezeigt werden, dass unter Berücksichtigung der spezifischen Varianz auch Testintelligenz und deduktives Schließen nur marginal über einen gemeinsamen Faktor beschrieben werden können.

Was lässt sich daraus schließen? Die Befunde verdeutlichen, dass die kognitiven Anforderungen, die sprach- und kulturfreiem logischen Schließen zugrundeliegen, von denen zu unterscheiden sind, die dem Testinstrument zum fächerübergreifenden Problemlösen zugrundeliegen. Leutner et. al. (2012) nehmen an, dass es sich bei fächerübergreifenden Problemlösekompetenzen um Kompetenzen der systematisch-strategischen und selbstregulierten Herangehensweise an komplexe Aufgabenstellungen ohne größeren Zeitdruck handelt, die bei Intelligenztestaufgaben in dieser Form nicht gefordert werden. Das heißt, es ist davon auszugehen, dass mit zunehmender Komplexität der Problemlöseaufgaben weitere Faktoren als ausschließlich *reasoning*-Aspekte virulent werden (u.a. Reeff, Zabal & Blech, 2006; Funke, 2003; Sternberg, 2003; Dörner, 1986). Die hier vorliegenden Befunde sind vor diesem Hintergrund zu erklären. Weiterhin bekräftigen sie, dass das Testinstrument einen eigenständigen Beitrag zur Klärung des fächerübergreifenden Kompetenzprofils der Studierenden leistet.

Werden die Befunde zu Disziplinunterschieden mitberücksichtigt, so konnte gezeigt werden, dass Teilkomponenten des fächerübergreifenden Problemlösens (bspw. Induktion, Korrelation, Fakten und Meinung) von Studierenden der Sozialwissenschaften besser gelöst werden können als von Studierenden der Geisteswissenschaften. Die Annahme der Studie war, dass sich diese Unterschiede zwischen den Disziplinen zeigen. Diese Annahme basiert darauf, dass in den Sozialwissenschaften der Umgang mit entsprechenden Problemstellungen ein relevanter Teil des Studiums ist. Das heißt, Daten und faktenbasierte Zusammenhänge müssen im Hinblick auf die Legitimität der Verknüpfung untersucht werden, Übergeneralisierungen sollten unter Berücksichtigung der Komplexität der Zusammenhänge vermieden werden. Vor dem Hintergrund des Experten-Novizen-Paradigmas sowie mit Blick auf Befunde zum analogen Problemlösen werden diese Ergebnisse noch diskutiert. Was bedeuten sie für die domänenspezifische Verortung des Konstruktes? Die Studie geht davon aus, dass fächerübergreifende Problemlösekompetenzen situations- und inhaltsübergreifend angewendet werden können und bei der Lösung neuartiger Probleme und komplexer Anforderungen erforderlich sind (Baumert et. al., 2000). Dabei liegt dem Testverfahren auch die Annahme zugrunde, dass Problemlösekompetenzen situiert erlernt werden und der Kontext beim Problemlösen nicht nur bei der Anwendung, sondern auch beim Erwerb eine bedeutsame Rolle spielt:

„Problem solving is merely a form of reasoning, that, [...] is deeply bound up with [...] the context in which it takes place.“ (Kirsh, 2009, S. 265)

Die Frage, inwieweit kognitive Komponenten des Problemlösens domänengebunden sind, ist allerdings bis heute nicht abschließend geklärt, die Ansätze der Informationsverarbeitung, der *situated cognition*, der Transfer- sowie der Expertiseforschung liefern Erklärungsansätze aus unterschiedlichen Blickwinkeln (u.a. Mayer, 2010; Mayer & Wittrock, 2006; Holyoack, 2005; Bassock, 2003; Barnett & Ceci, 2002; Nickerson, 1998; Salomon & Perkins, 1989). Die Befunde zu Disziplinunterschieden legen nahe, dass domänenspezifisch vermittelte Expertise bei der Lösung fächerübergreifender Problemstellungen, in Abhängigkeit der kognitiven Anforderung eine Rolle spielt.

18. Fächerübergreifende Problemlösekompetenzen bei Studierenden

Neben der Konstruktion des Testverfahrens und der Exploration der Struktur des Konstruktes sollte vor allem die Leistung der Studierenden im fächerübergreifenden Problemlösen beschrieben werden, über den Einbezug von Kovariaten sollten Unterschiede zwischen Studierenden und Bedingungen fächerübergreifender Problemlösekompetenzen erklärt werden.

18.1 Fächerübergreifende Problemlösekompetenzen bei Studierenden und die Rolle der Expertise

Die Zusammenhänge zwischen Expertise und Problemlösen wurden über Vergleiche zwischen Experten und Novizen bereits zahlreich untersucht und im Experten-Novizen-Paradigma zusammengefasst (bspw. Chi, 2006; Novick & Bassok, 2005; Chi, & Friedman, 2003; Ericsson, 2003b; Sternberg, 2003; Catrambone, 1996; Sternberg & Frensch, 1992; Novick, 1988; Duncker, 1945). Dieses geht davon aus, dass die Entwicklung professioneller Kompetenzen ein sukzessiver Aufbau von Fähigkeiten und Wissen ist. Expertise nimmt dabei auf Problemlösekompetenzen in unterschiedlicher Weise Einfluss, so bspw. auf die *task performance*, das heißt, die Testleistung, Strategienanwendung oder auch Problemrepräsentationen, auf die Struktur der Wissensbasis, auf *chunking* oder auch auf die Wissensorganisation. Im Rahmen dieser Studie wurde Expertise als Studienerfahrung über die Anzahl der Semester sowie die Studiendisziplin definiert. Die Befunde zeigen durchaus Unterschiede beim fächerübergreifenden Problemlösen in Abhängigkeit der Expertise.

Was ließ sich genau feststellen? Es konnte beobachtet werden, dass Expertise bei der Lösung von auf Wahrscheinlichkeiten basierenden Problemen, beim Unterscheiden von subjektiver und objektiver Information sowie bei der Zerlegung komplexer Probleme, das heißt, beim Erkennen multipler Einschränkungen, relevant wird. Weiter lassen die Befunde darauf schließen, dass die Bekanntheit der Kontextualisierung der Problemstellung („Handwerker“ vs. „Migration“ beim induktiven Schließen),

die Logik der Verknüpfung („Alkohol“ vs. „Neuwagen“ beim korrelativen Schließen) sowie die Schwierigkeit der Informationsdarstellung („Arbeitslose“ vs. „Raucher“ beim Unterscheiden von Fakten und Meinung) in Zusammenhang mit Expertise zu stehen scheint. Die Modalität der Informationsdarstellung (Zielevaluation) steht hingegen nicht in Zusammenhang mit Expertise. Wird die über die Disziplin erworbene Expertise mit der über das Semester erworbenen Expertise verglichen, so zeigen die Befunde, dass zwischen beiden differenziert werden muss. Es war zu beobachten, dass die Abschätzung von Wahrscheinlichkeiten für geisteswissenschaftlich orientierte Studierende eine Schwierigkeit darstellte (induktives Schließen), sich jedoch die Semester nicht bedeutsam unterschieden. Erfordert die Problemlösung jedoch zusätzlich die Berücksichtigung der Plausibilität des logischen Schlusses, bzw. die Abschätzung der Intensität des Zusammenhangs zweier Variablen, scheint auch die Erfahrung eines höheren Semesters erforderlich zu werden. Das heißt, mit zunehmendem Semester differenzieren die Studierenden die Fähigkeit aus, illegitime kausale Schlüsse zu vermeiden, auch wenn die gegebenen Informationen einen Zusammenhang nahelegen.

Das Semester nahm keinen Einfluss auf die Fähigkeit zur Handlungsplanung. Erstsemester haben ebenso Schwierigkeiten wie Dritt- und Sechstsemester. Hier konnte lediglich der Einfluss der Studiendisziplin festgestellt werden.

Wie können diese Befunde erklärt werden? Orientiert am Experte-Novizen-Paradigma kann daraus geschlossen werden, dass sich die Studienerfahrung, also hier die Auseinandersetzung mit Problemstellungen im Rahmen des sozialwissenschaftlich ausgerichteten Studiums, bzw. die Intensität und Häufigkeit der Konfrontation mit entsprechenden Problemstellungen über mehrere Semester hinweg darauf auszuwirken scheint, dass illegitime Schlüsse besser erkannt, bzw. mögliche *constraints* besser berücksichtigt werden können. Dies ist kongruent mit bisherigen Befunden (bspw. Novik & Bassok, 2005; Catrambone, 1996; Schiano, Cooper, Glaser & Zang, 1989; Voss, Greene, Post & Penner, 1983; Chi, Glaser & Rees, 1982; Chi, Feltovich & Glaser, 1981; Simon & Simon, 1978).

Ausgehend von Befunden der Transferforschung lässt sich argumentieren, dass Studierende höherer Semester, bzw. Studierende der sozialwissenschaftlichen Disziplin, die jeweiligen Problemlöseaufgaben deshalb besser lösen konnten, da sie Struktur(un)ähnlichkeiten zwischen vorliegenden und bereits im Studium erarbeiteten Problemstellungen besser abgleichen konnten (u.a. Nokes, 2009; Bassok, 2003; Holyoack, & Koh, 1987; Gick & Holyoack, 1983; Genter, 1982, 1983; Reed et al. 1974). Expertise wäre dementsprechend ein besserer Umgang sowohl mit Oberflächen-, als auch mit Tiefenstrukturmerkmalen. Daraus folgt, dass zwischen den Studierenden auch Unterschiede im *framing* bestehen könnten. Erkennen Sozialwissenschaftsstudierende Merkmale des Problems, bzw. seiner Struktur besser und identifizieren relevante Informationen, bspw. Einschränkungen besser, schätzen sie Wahrscheinlichkeiten besser ein? Dies würde sich auch darauf auswirken, welcher Suchraum zur Problemlösung aktiviert und welcher Pfad der Problemlösung gewählt wird. Es lässt sich bislang nur vermuten, dass im Rahmen sozialwissenschaftlicher Studiengänge der Umgang mit

auf Wahrscheinlichkeitsschätzungen basierenden Problemstellungen, komplexen Problemstellungen, die das Erkennen multipler Einschränkungen erfordern, in höherem Maß als in geisteswissenschaftlichen Studiengängen eine Rolle spielen. Dass bei den Studierenden der vorliegenden Studie Expertise erworben wurde und zu einer besseren Leistung führt, konnte gezeigt werden. Dabei zeigen die Befunde auch, dass sich die Expertise aufgrund eines höheren Semesters und die Expertise aufgrund eines anderen Studienfaches unterscheiden. Welche Form der Expertise jeweils erworben wurde und welche Fähigkeiten den Studierenden hilfreich waren, konnte im Rahmen der Studie nicht geklärt werden.

18.2 Weitere Schwierigkeiten beim fächerübergreifenden Problemlösen

Für eine Prognose fächerübergreifender Problemlösekompetenzen ist die Identifikation von Schwierigkeiten zentral (u.a. Mayer, 2010; Pretz, Naples & Sternberg, 2004; Funke, 2003; Jonassen, 2000; Kotovsky & Simon, 1990; Greeno & Simon, 1988). Neben den bisher beschriebenen Schwierigkeiten beim Problemlösen wurden weitere Merkmale identifiziert, die Einfluss auf die Schwierigkeit nehmen. Im Rahmen der vorliegenden Studie zeigt der Blick auf deduktives Schließen sowie die Fähigkeit zur Handlungsplanung, dass einerseits die Anzahl der zu manipulierenden Variablen ein Ansatzpunkt zur Modifikation der Schwierigkeit war. Die Verknüpfung mehrerer Variablen beim deduktiven Schließen war mit großen Schwierigkeiten verbunden. Bei der Handlungsplanung ließ sich darüber hinaus feststellen, dass nicht nur die Fähigkeit zur Verknüpfung, sondern auch die Fähigkeit zur Verknüpfung mehrerer Variablen vor dem Hintergrund multipler Einschränkungen noch größere Schwierigkeiten bereitet. Dabei konnten die Studierenden nach der Anzahl der Einschränkungen, die sie erkennen können, unterschieden werden. Der Blick auf die Fähigkeiten zur Zielevaluation konnte zeigen, dass des Weiteren die Modalität der Darbietung (Text, Grafik oder Statistik) das Anforderungsniveau beeinflusste, wobei textbasiert gegebene Informationen leichter identifiziert werden konnten als Informationen, die Statistiken entnommen werden mussten. Aufgrund dieses Befundes ist auch bei den Testitems zum deduktiven Schließen nicht auszuschließen, dass die Darstellung der Information (Statistiken beim Testitem „Gesundheitskosten“) einen zusätzlichen erschwerenden Aspekt darstellte. Insgesamt konnten damit bisherige Befunde zur Manipulation der Problemschwierigkeit repliziert werden (u.a. Fleischer, Wirth, Rumann & Leutner, 2010; Jonassen, 2000). Weiter konnte gezeigt werden, dass auch innerhalb geschlossener Probleme (*well-defined problems*) Ansatzpunkte zur Variation der Schwierigkeit herangezogen werden können, um komplexeres planerisches und strategisches Vorgehen zu simulieren. Zur Beschreibung und Prognose von Schwierigkeiten beim fächerübergreifenden Problemlösen steht damit nicht nur die Anzahl der zu manipulierenden Variablen, sondern auch die Art der kognitiven Operation und die Darstellungsmodi der Variablen zur Verfügung.

18.3 Die Rolle subjektiver Überzeugungen (I): Epistemologische Überzeugungen und Selbstwirksamkeit

Inwieweit können die Unterschiede beim fächerübergreifenden Problemlösen näher beschrieben werden? Konzeptualisiert werden epistemologische Überzeugungen mit den Dimensionen der *Sicherheit des Wissens* („Wissen ist sicher und steht ein für allemal fest“ ggü. „Wissen ist unsicher und provisorisch“), *Einfachheit des Wissens* („Wissen besteht aus isolierten Fakten“ ggü. „Wissen besteht aus miteinander vernetzten Sachverhalten und Begriffen“), *Quelle des Wissens* („Wissen wird von Autoritäten oder Experten weitergegeben“ ggü. „Wissen ist das Ergebnis selbständiger Konstruktionsprozesse“) und *Rechtfertigung* („Wissen kann unbesehen geglaubt werden“ ggü. „Wissen bedarf der Rechtfertigung durch Argumente, empirische Daten etc.“). Als Kernkomponente epistemologischer Überzeugungen wird die Überzeugung der Sicherheit des Wissens begriffen. In Bezug auf alle Dimensionen wird die jeweils unreflektiertere Ausprägung als „naiv“ bezeichnet, die reflektiertere als „reif“ (Hofer, 2000; Schommer, 1990). Ihre Relevanz erhalten die epistemologischen Überzeugungen darüber, dass sie als Komponenten informellen Wissens Lernprozesse in vielfältiger Weise beeinflussen (u.a. Bromme, Pieschl & Stahl, 2010; Buehl, 2008; Cano & Cardell-Elawar, 2008; Hofer, 2008; Moschner, Gruber & die Studienstiftungsgruppe EPI, 2005; Bendixen & Rule, 2004; Urhahne & Hopf, 2004; Baumert et. al., 2000; Hofer, 2000; Schommer, 1992, 1990).

Wird in erster Linie die Frage nach Selektionseffekten bei den Studierenden gestellt – bislang liegen diesbezüglich uneinheitliche Befunde vor (u.a. Trautwein & Lüdtke, 2008; Becher & Trowler, 2000; Ylijoki, 2000; Lindblom-Ylänne, 1999) – so verweisen die Ergebnisse vorliegender Studie auf den Einfluss der Studiensozialisation. Die Befunde zeigen zwar, dass Sozialwissenschaftsstudierende eher als Geisteswissenschaftsstudierende davon ausgehen, dass sich Wissen weiterentwickelt und nicht statisch ist. Selektionseffekte konnten allerdings nicht beobachtet werden, Unterschiede zwischen den Disziplinen waren nicht im ersten, sondern lediglich im dritten Semester beobachtbar. Dies lässt die Annahme von Sozialisationseffekten zu, wobei sich einerseits die Überzeugungen mit zunehmender Studienexpertise ausdifferenzieren, andererseits die Unterschiede zwischen den Disziplinen in Bezug auf die Natur des Wissens und der Wissensentwicklung bis zum Studienende zu nivellieren scheinen. Signifikante Unterschiede bei der Überzeugung von der Transferierbarkeit fachspezifischer Problemlösemethoden auf neue Probleme verweisen allerdings darauf, dass durchaus Sozialisationsunterschiede je nach Disziplinen auch bis zum Studienende hin bestehen, diese jedoch scheinbar in Abhängigkeit von der Art der subjektiven Überzeugung. Die Befunde zeigen damit, dass eine Analyse des inhärenten *habitus* der Disziplinen, wie auch von Hofer (2000) angeregt, und die Untersuchung seines Einflusses auf die Entwicklung subjektiver Überzeugungen bei Studierenden aufschlussreich wäre (auch Kaartinen-Koutaniemi & Lindblom-Ylänne, 2008). Die vorliegenden Befunde zeigen dabei auch, dass die größten Effekte zwischen dem ersten und dritten Semester vorliegen. Dies steht in Übereinstimmung mit bereits vorhandenen Daten (Trautwein & Lüdtke, 2008).

Insgesamt zeigen die Befunde der Studie, dass sich die epistemologischen Überzeugungen der Studierenden zu „entwickeln“ scheinen – wobei der Terminus der „Weiterentwicklung“ unter dem Vorbehalt des Querschnittsdesigns zu sehen ist. Dahingehend zeigen kleine Effekte, dass die Expertise Einfluss nimmt, wobei zwischen Expertise im Sinne der Studiendisziplin und Expertise im Sinne des Semesterverlaufs zu unterscheiden ist. Das bedeutet, Sozialwissenschaftsstudierende (Sicherheit) und höhere Semester (Sicherheit, Quelle) haben reflektiertere epistemologische Überzeugungen als Studierende der Geisteswissenschaften und jüngere Semester. Studierende höherer Semester gehen eher davon aus, dass sich Wissen weiterentwickelt, flexibel ist und auch über die eigene Auseinandersetzung konstruiert werden kann und nicht nur von Autoritäten, bspw. Dozenten oder Lehrbüchern weitergegeben wird. Sozialwissenschaftsstudierende gehen eher davon aus, dass sich Wissen weiterentwickelt. Die Auseinandersetzung mit Wissen und der Weiterentwicklung von Wissen über die Studiensemester hinweg sowie in Abhängigkeit der Disziplin scheint sich damit auf die Entwicklung reflektierterer epistemologischer Überzeugungen auszuwirken. Diese Ergebnisse replizieren Befunde, die bereits zeigen konnten, dass Studierende mit zunehmender Studienexpertise reflektiertere Annahmen im Hinblick auf Wissen und Wissensentwicklung entwickeln und sich die epistemologischen Überzeugungen je nach Disziplin (sogenannte harte vs. weiche Fächer) unterscheiden (u.a. Hofer, 2008; Kaartinen-Koutaniemi & Lindblom-Ylänne, 2008; Trautwein & Lüdtke, 2008; Palmer & Marra, 2004; Perry, 1999).

Inwieweit stehen nun epistemologische Überzeugungen und fächerübergreifende Problemlösekompetenzen in Zusammenhang? Die Ergebnisse zeigen, dass Studierende, die von der Veränderbarkeit von Wissen überzeugt sind, signifikant bessere Testleistungen erbringen. Eine differenzierte Analyse konnte weiter zeigen, dass die Überzeugung von der Sicherheit des Wissens insbesondere diejenigen Testleistungen beeinflusste, die auf induktivem und korrelativem Schließen basieren. Ebenso beeinflusst die Überzeugung von der Veränderlichkeit des Wissens in kleinem Maße die Fähigkeit zur Handlungsplanung. Der Vergleich weniger reflektierter und reflektierterer Studierender stellte den Effekt der epistemologischen Überzeugungen auf induktives und korrelatives Schließen nochmals heraus. Das bedeutet, dass Studierende, die ihr fachspezifisch erlerntes Wissen als veränderbar ansehen diejenigen Probleme signifikant besser lösen konnten die auf Unsicherheiten basierten und auf der Evaluation von Wahrscheinlichkeiten. Wie können diese Befunde erklärt werden? Die Ergebnisse unterstreichen die Rolle epistemologischer Überzeugungen als Leistungsprädiktoren (u.a. Hofer, 2000), wobei sie den Effekt domänenspezifischer epistemologischer Überzeugungen auf fächerübergreifende, also neue und unbekannte Probleme erweitern. Allerdings spielt nur diejenige Überzeugung eine Rolle, die sich auf die Natur des Wissens bezieht. Die Überzeugung davon, wie Wissen entwickelt wird, beeinflusste nicht direkt die Problemlösekompetenz. Sie konnte jedoch im Rahmen der Clusteranalyse als Variable identifiziert werden, die ebenfalls zwischen Studierendengruppen mit unterschiedlich ausgeprägten fächerübergreifenden Problemlösekompetenzen unterscheidet.

Wie kann der Einfluss epistemologischer *beliefs* auf den Umgang mit unsicherer Information erklärt werden? Induktives und korrelatives Schließen erfordern die Reflexion fehlerhafter Informationen und Annahmen. In dieser Hinsicht konnten Studien zeigen, dass Studierende mit weniger reflektierten Überzeugungen eher Behauptungen aufstellen, die in ihrer Einseitigkeit nicht mit den Daten übereinstimmen (Quain & Alvermann, 1995; Schommer, 1990). Weiterhin konnte gezeigt werden, dass Studierende, die davon ausgingen, dass Wissen aus kontroversen, verknüpften Ideen und nicht isolierten Fakten besteht, Texte besser verstehen, die unterschiedliche Standpunkte gegenüberstellten (Kardash & Scholes, 1996). Die Ergebnisse der vorliegenden Studie ergänzen diese Befunde und zeigen, dass reflektiertere Überzeugungen, insbesondere die Überzeugung von der Veränderbarkeit des Wissens Problemlöser befähigt, mit denjenigen Testaufgaben besser umzugehen, die auf Wahrscheinlichkeitsschätzungen basieren und unerklärte, bzw. unvollständige Daten beinhalten.

Daneben wurde auch die Frage beleuchtet, unter welchen Umständen epistemologische Überzeugungen wirksam werden. Dass Selbstwirksamkeit in Zusammenhang mit epistemologischen Überzeugungen zu stehen scheint, wurde bereits angenommen (Schommer-Aikins et. al. 2005; Bendixen & Rule, 2004). Ungeklärt ist allerdings die Frage, wie die Zusammenhänge beschrieben werden können. Die bisherige Annahme ist, dass Selbstwirksamkeit die Entwicklung reflektierter *beliefs* beeinflusst und nicht nur bei der Anwendung von Problemlösestrategien und -methoden relevant wird, sondern insbesondere auch bei der Überzeugung, eigene Problemlösemethoden selbsttätig verändern zu können (Schommer-Aikins et. al., 2005; Bendixen & Rule, 2004).

Die vorliegende Studie konnte diese Annahme empirisch stützen und konnte zeigen, dass reflektiertere epistemologische Überzeugungen mit einer höheren Selbstwirksamkeit einhergehen, bzw. umgekehrt. Sie zeigt, dass Studierende mit einer höheren Selbstwirksamkeit eher davon ausgehen, dass Wissen veränderbar und flexibel ist. Ebenso gehen sie eher davon aus, dass Wissen auch ein Ergebnis individueller Konstruktionsprozesse sein kann. Weiter verdeutlichen die Befunde, dass die Überzeugung von der Wirksamkeit des eigenen Handelns ein relevanter Faktor bei der Erklärung der Ausprägung der Überzeugung im Hinblick auf die Veränderbarkeit des Wissens ist. Dies kann darüber erklärt werden, dass die individuelle Person und deren Auseinandersetzung mit Wissen im Zentrum des Konzeptes epistemologischer Überzeugungen steht (Hofer, 2008). Eine Voraussetzung für diese Auseinandersetzung mit Wissen scheint das Vertrauen in die eigene Handlungsfähigkeit und -wirksamkeit zu sein. Die Befunde stützen damit die Annahme von Bendixen und Rule (2004), die davon ausgehen, dass eine höhere Selbstwirksamkeit Studierende zu mehr Selbstsicherheit ermutigen kann, kognitive Werkzeuge zu nutzen, eigene Ideen zu integrieren und unterschiedliche Konzepte zusammenzubringen. Strike und Posner konnten diese Annahme bereits empirisch prüfen (1992). Trotz der vorliegenden Befunde bleibt weiter ungelöst, wie epistemologische Überzeugungen und die Überzeugung von der Wirksamkeit des Handelns in Zusammenhang stehen. Ermöglicht eine höhere Selbstwirksamkeit eine tiefere Auseinandersetzung mit der Veränderbarkeit des Wissens? Oder erleichtert die Überzeugung von der Komplexität und Veränderbarkeit des Wissens bspw. den

Umgang mit schwierigen Problemen und führt deshalb zu einer höheren Selbstwirksamkeit? Zu vermuten bleibt weiterhin auch, dass die Überzeugung eine Rolle spielt, Wissen auch eigentätig verändern zu können, das heißt, neue eigene Gedanken und Ideen eigentätig entwickeln zu können. Die Erfassung epistemologischer *beliefs* in der hier vorliegenden Form erhebt lediglich die Überzeugung von der Veränderbarkeit des Wissens und nicht inwieweit es sich die Studierenden selbst zutrauen, am Prozess der Veränderung des Wissens auch teilnehmen zu können. Aufschlussreich wäre letzten Endes auch die Frage, inwieweit Studierende mit höherer Selbstwirksamkeitserwartung und reflektierteren *beliefs* die Auseinandersetzung mit Wissen auch für wichtiger halten als Studierende mit geringerer Selbstwirksamkeitserwartung. Dies ist deshalb interessant, da subjektive Überzeugungen zwar bewirken, dass Vorstellungen davon gebildet werden, wie Situationen verändert werden und inwieweit man sich selbst dazu in der Lage sieht, diese Veränderungen bewirken zu können. Die Voraussetzung hierfür ist allerdings, dass die jeweilige Situationsveränderung auch als wünschenswert eingeschätzt wird (Schwarzer, 1995).

18.4 Die Rolle subjektiver Überzeugungen (II): Transfer

Die Evaluation der Überzeugung der Studierenden, dass in der Fachdomäne erworbene Problemlösemethoden auf neue, fächerübergreifende Probleme angewendet werden können, basierte darauf, dass der Transfer im Zentrum des Konzeptes der Schlüsselkompetenzen steht. Das Konzept der Schlüsselkompetenzen geht davon aus, dass fächerübergreifende Kompetenzen zwar situations- und inhaltbezogen gelernt, allerdings situations- und inhaltunabhängig in neuen und unbekanntem Situationen angewendet werden können. Bei fächerübergreifenden Problemlösekompetenzen, so die Annahme der Studie, spielt dementsprechend auch die Überzeugung eine Rolle, die fachspezifisch erlernten Problemlösekompetenzen auch auf andere Probleme und Kontexte übertragen und anwenden zu können. Aus den berichteten Befunden zum Einfluss subjektiver Überzeugungen auf Handeln, Lernen und Leistung wurde abgeleitet, dass Studierende, die nicht von der Transferierbarkeit überzeugt sind, beim Erlernen fächerübergreifender Problemlösekompetenzen weniger motiviert sind, da sie nicht die Notwendigkeit sehen, dementsprechend weniger Üben und Ausprobieren und schließlich beim Lösen fächerübergreifender Probleme schwächer sind als Studierende mit einer höheren Überzeugung. Die Befunde der vorliegenden Studie zeigen allerdings keinen Zusammenhang zwischen der Transferüberzeugung und der Problemlöseleistung. Signifikante Korrelationen sind lediglich mit Studieninteresse und Selbstwirksamkeit zu beobachten. Das bedeutet, dass Studierende mit einem höheren Fachstudieninteresse und Studierende mit einer höheren Selbstwirksamkeitserwartung von der Anwendbarkeit fachspezifisch erlernter Arbeitsmethoden auf unbekannte Probleme überzeugt sind. Die Überzeugung, dass das eigene Handeln weniger wirksam ist sowie ein geringeres Studieninteresse gehen mit einer geringeren Transferüberzeugung einher. Daneben ließ sich auch feststellen, dass die Studiendisziplin Einfluss auf die Überzeugung von der

Transferierbarkeit fachspezifischer Problemlösemethoden nimmt. Es zeigte sich, dass Studierende der Sozialwissenschaften überzeugter sind, dass ihre Problemlösemethoden nicht nur fachspezifisch anzuwenden sind. Interessanterweise nahm die Überzeugung vom Transfer fachspezifischer Problemlösemethoden auf neue, fächerübergreifende Probleme im Laufe der Semester ab, Erstsemester sind signifikant überzeugter als Sechstsemester.

Die Befunde zeigen damit insgesamt, dass zwar Studierende mit höherer Selbstwirksamkeitserwartung überzeugter sind, ihre im Fach erworbenen Problemlösemethoden auch außerhalb ihrer Studiendisziplin anwenden zu können. Auch ein höheres Studieninteresse führt dazu, dass Studierende eher vom Transfer überzeugt sind. Im Laufe des Semesters geht diese Überzeugung jedoch zurück. Dies ist insbesondere deshalb ein bedeutsamer Befund, da die vertiefte Auseinandersetzung mit der Fachwissenschaft im Laufe des Semesters eigentlich zu einer gegenteiligen Entwicklung führen sollte (u.a. Bransford & Schwarz, 1999; Mayer & Wittrock, 1996; Salomon & Perkins, 1989): „The belief in transfer lies at the heart of our educational system. Most educators want learning activities to have positive effects that extend beyond the exact conditions of initial learning“. (Bransford & Schwarz, 1999, S. 61)

18.5 Studieninteresse

Der Einfluss emotionaler Aspekte auf Problemlösen wird vor allem über Verarbeitungsmodi und die Aufmerksamkeit gegenüber relevanten Problemstrukturen gesteuert (De Sousa, 2008; Schwarz & Skurnik, 2003). Emotionale Einflüsse auf Problemlösen werden deshalb auch als *cognitive tuning* bezeichnet (Schwarz & Skurnik, 2003; Martin et al., 1993). In Bezug auf den Zusammenhang emotionaler Aspekte und Leistung wird neben Selbstwirksamkeit das Interesse besonders bedeutsam (Schiefele, Krapp und Winteler, 1992; Mayer, 1998). Ausgehend davon, dass Interesse mit einer Korrelation von ungefähr $r=.30$ zwischen Interesse an einem Schulfach und der in diesem Fach erzielten Leistung Einfluss nimmt (Renninger, Hidi & Krapp, 1992; Schiefele, Krapp & Winteler 1992; Schiefele, Krapp & Schreyer, 1993), wurde erwartet, dass auch für die vorliegende Studie entsprechende Zusammenhänge zu finden sind. Diese Vermutung ließ sich jedoch nicht bestätigen. Auffällig war darüber hinaus, dass das Studieninteresse mit zunehmendem Semester signifikant abnahm.

Wie sind diese Befunde zu erklären?

In Bezug darauf, dass keine Zusammenhänge zwischen Leistung und Interesse gefunden werden konnten, lassen sich unterschiedliche Erklärungsansätze heranziehen: Der Zusammenhang zwischen Interesse und Leistungsergebnissen wird oftmals über die Verarbeitungstiefe und -breite oder qualitative Wissensunterschiede beschrieben (Schiefele, 1992, 1988; Schiefele, Krapp, Wild & Winteler 1993; Wild, Krapp & Winteler 1992). Verarbeitungsbezogene Aspekte wie bspw. Heuristiken des Problemlösens, Tiefe des Problemverständnisses oder auch die Reflexion des zur

Lösung des Problems erforderlichen Schlusses wurden im Rahmen dieser Studie nicht erfasst, die quantitative Erfassung der Kompetenz im Sinne der Messung über eine richtig-falsch-Zuordnung greift deshalb gegebenenfalls zu kurz. Daneben kann jedoch auch vermutet werden, dass die Operationalisierung des Interessenskonstruktes zu unspezifisch war: So zeigen Befunde, dass insbesondere kognitives Interesse mehr als emotionales Interesse Einfluss auf Lernen und Leistung nimmt (Harp & Mayer, 1997, 1998; Kintsch 1980, nach Harp & Mayer, 1997). Des Weiteren verweisen Metaanalysen darauf, dass die subjektiv eingeschätzte Passung zum Studium höher mit Leistung korreliert als das global abgefragte Studieninteresse (Mayr, 2011; Schiefele, Krapp & Schreyer, 1993; Krapp & Schreyer, 1994). Aus den hier vorliegenden Befunden kann dementsprechend die Vermutung abgeleitet werden, dass der Zusammenhang zwischen Interesse und Leistung gegebenenfalls eher über spezifische Interessensfaktoren wie beispielsweise selbstbestimmte Motivation oder Autonomieerleben besteht (Mayr, 2011; Deci & Ryan, 1987; 2000; De Charms, 1968). Auch Krapp und Ryan (2002) betonen, dass es gerade im pädagogischen Kontext nicht nur darauf ankommt, wie stark jemand motiviert ist, sondern vielmehr von welcher „Qualität“ die Motivation ist. Auch der hier vorliegende Befund, dass das Studieninteresse vom ersten bis zum sechsten Semester geringer wird, könnte mit einem gegebenenfalls zu unspezifisch erfassten Studieninteresse begründet werden. Zwar liegen ähnliche Befunde eines geringer werdenden Interesses am Studium in höheren Semestern bereits vor (Niketka, 2006), es könnte jedoch auch vermutet werden, dass Studieninteresse in höheren Semestern aus anderen Facetten besteht als in niedrigeren Semestern.

18.6 Kompetenzstrukturen

Zum Abschluss der Studie wurde geprüft, inwieweit Zusammenhänge zwischen Teilkompetenzen nachgezeichnet werden können. Dem Modell liegt die Annahme zugrunde, dass die beobachteten Teilkompetenzen fächerübergreifenden Problemlösens nicht nur als Einzelkomponenten wichtig werden. Vielmehr werden sie in ihrem Zusammenwirken notwendig, um ein Problem erfolgreich bearbeiten zu können. In Anlehnung an Weinert (1999) wurden die Kompetenzen als funktionale kognitive Leistungsdispositionen definiert, die sich auf bestimmte Klassen von Situationen und Anforderungen beziehen. Das Strukturmodell konnte zeigen, dass unterschiedliche kognitive Anforderungen in Zusammenhang gebracht werden können. Insbesondere das Erkennen von Einschränkungen, das heißt, das Erkennen von Einschränkungen auf Basis einer Evaluation der vorliegenden Informationen – sei es ein Abwägen der Intensität der Zusammenhänge beim korrelativen Schließen oder ein Vergleich subjektiver und objektiver Informationen oder multipler Einschränkungen – stehen in Zusammenhang (Teilkompetenzen der Situationsanalyse: induktives und korrelatives Schließen; Teilkompetenzen der Situationsanalyse und Handlungsplanung: korrelatives

Schließen, Fakten und Meinung unterscheiden und Handlungsplanung; bewertendes und planendes Schließen im Rahmen der Zielevaluation und Handlungsplanung).

Die Clusteranalysen integrierten abschließend kognitive, metakognitive und motivationale Faktoren. Dabei konnte gezeigt werden, dass es bei der Unterscheidung zwischen Studierenden bedeutsam ist, sowohl die fächerübergreifende Problemlösekompetenz als auch die epistemologische Überzeugung und die Selbstwirksamkeit zu berücksichtigen.

19. Schlussfolgerungen für die Lehre und Ableitungen für die Analyse von *academic skills*

In der Einleitung der Arbeit wurde dargestellt, dass Studierende im beruflichen Alltag mit Problemstellungen umgehen müssen, die neuartig und komplex sind und den Transfer domänenspezifischer Kompetenzen erfordern. Um diese lösen zu können, müssen nicht nur Informationen miteinander verknüpft werden. Vielmehr müssen Studierende die Gültigkeit von Informationen abwägen, mit unvollständigen Daten umgehen, Entscheidungen auf Basis einer unsicheren Datenlage treffen oder auch komplexe Einschränkungen erkennen. Die Studie konnte verschiedene Schwierigkeiten beim Umgang mit entsprechenden Anforderungen aufdecken. Die Frage ist, was dies für die Lehre bedeutet. Im Theorieteil wurden bereits Ansätze zur Überwindung trägen Wissens erläutert. Dabei stand die Frage im Zentrum, welche Bedingungen positiven Transfer begünstigen. Was lässt sich vor diesem Hintergrund aus den hier vorliegenden Befunden für die Lehre ableiten?

Die Notwendigkeit der Lehre von Transfer wird damit begründet, Studierenden zu vermitteln, Kompetenzen und Wissen aus unterschiedlichen Disziplinen für eine neue Problemlösung heranzuziehen, domänenspezifisch erworbenes Wissen auch in neuen Kontexten anzuwenden und flexibles Wissen zu erwerben. Die Auseinandersetzung mit einem konkreten Problem soll dessen Lösung auf Basis einer argumentativen Auseinandersetzung fördern. Der Lehransatz des *Problem based learning* bspw. setzt sich einerseits die Überwindung trägen Wissens und die Auseinandersetzung mit Strategien des Problemlösens zum Ziel, andererseits auch die Förderung von *reasoning*-Strategien. Die hier vorliegenden Befunde verdeutlichen, dass im Weiteren insbesondere auch die Auseinandersetzung mit den jeweils gegebenen Problemelementen neben der Vermittlung und Erprobung von Problemlösestrategien eine Rolle spielt, um zu einer erfolgreichen Problembearbeitung gelangen zu können: Welche Einschränkungen sind probleminhärent gegeben? Wie ist die Validität der gegebenen Information einzuschätzen? In welchem Kontext steht die gegebene Information? Wie legitim ist die Verknüpfung der gegebenen Informationen? Ein konzeptuelles Verständnis des Problems und die Berücksichtigung der Tiefenstrukturen der Problemstellung oder auch das Ignorieren irrelevanter Informationen stellen Voraussetzungen für die Problemlösung dar (auch Catrambone, 1996; Novick, 1988).

Daneben zeigen die Befunde, dass der Blick ebenso auf motivationale und metakognitive Faktoren gerichtet werden muss. Im Zusammenhang mit der Förderung kritischen Denkens beschrieb Halpern (1998) bereits ein Modell, um metakognitive Kompetenzen zu fördern. Für die Umsetzung ihres Modells verwies sie darauf, dass sich im Zusammenhang mit der Förderung kritischen Denkens die Sichtweise auf Lehre verändern müsse, wobei der Fokus auf den Transfer gelegt werden sollte und der Aspekt des Transfers der Ausgangspunkt für Lehre sein sollte:

„In critical thinking instruction, the goal is to promote the learning of transcontextual thinking skills and the awareness of and the ability to direct one's own thinking and learning. Although thinking always occurs within a domain of knowledge, the usual methods that are used for teaching content matter are not optimal for teaching the thinking skills [...] because instruction [...] focuses on content knowledge [...] instead of the transferability [...]. Fortunately, there already are powerful models of human learning that can be used as a guide for the redesign of education [...].“ (Halpern, 1998, S. 451)

Die Befunde zeigen, dass der Fokus auf den Aspekt des Transfers vor allem auch deshalb besonders wichtig ist, da die Studierenden im Rahmen der vorliegenden Studie mit höherem Semester immer weniger davon ausgingen, dass ihnen fachspezifisch erlernte Problemlösemethoden auch in anderen als den im Studium bearbeiteten Problemstellungen behilflich sein können. Der Förderung subjektiver Überzeugungen räumt Halpern in ihrem Modell als exekutiven Funktionen einen zentralen Platz ein. Lehrstrategien wie Selbstinstruktion, die Überwachung eigener Lösungsschritte oder des eigenen Umgangs mit Wissen, der Aufbau eines systematischen Wissens über subjektive Überzeugungen sollten die Schärfung der metakognitiven Sensitivität der Studierenden zum Ziel haben. Obwohl viele Fragen mit Blick auf die Mechanismen epistemologischer Überzeugungen und insbesondere auch der Rolle domänenspezifischen Wissens offen sind (Kienhues, Bromme & Stahl, 2008), gibt es bereits Modelle, um die epistemologischen Überzeugungen der Studierenden zu fördern. Bendixen und Rule (2004) schlagen ein dreischnittiges Modell vor, in dessen Zentrum die Auseinandersetzung mit der Plausibilität epistemologischer Überzeugungen auf jeder Stufe ihrer Entwicklung steht. Zweifel an den Überzeugungen im Hinblick auf die Natur des Wissens und die Wissensentwicklung, die Akzeptanz der eigenen Verantwortlichkeit für die subjektiven Überzeugungen ebenso wie das Erkennen von Strategien zur Veränderung der Überzeugungen – bspw. Reflexion oder Auseinandersetzung mit anderen Studierenden – nennen sie als Lehrmethoden. Für die Lehre müssen die zur Verfügung gestellten Probleme verständlich, kohärent und plausibel, motivierend und adaptiv sein, um Zweifel bei den Studierenden anzuregen. Weiterhin sollten Studierende die Möglichkeit erhalten, Problemlösungen zu hinterfragen und kritisch die Relevanz subjektiver Überzeugungen bei der Auseinandersetzung mit Problemen zu diskutieren.

Schließlich konnte im Rahmen der Studie ebenso nachgezeichnet werden, dass der kritische Umgang mit Wissen mit der Einschätzung der Wirksamkeit des eigenen Handelns zusammenhängt. Das

Erleben kognitiver Konflikte ist deshalb ein Lehrziel. Ein weiteres muss sein, die Einschätzung der Wirksamkeit des eigenen Handelns der Studierenden zu stärken und Studierenden das Erleben eigener Kompetenz zu ermöglichen.

Für die Beschreibung der im Studium zu vermittelnden *academic skills* bedeutet dies, dass die bislang deskriptiv beschriebenen Kompetenzbereiche sorgfältig operationalisiert werden müssen, um sie nicht nur empirisch zu überprüfen, sondern insbesondere auch bedeutsame Leistungsprädiktoren zu identifizieren und Wirkungszusammenhänge zu beschreiben. Dabei ist weiterhin zu berücksichtigen, dass nicht nur direkt zu beobachtende Kompetenzen beschrieben werden sollten, sondern insbesondere auch *skills* wie bspw. epistemologische Überzeugungen, deren Effekte nicht direkt beobachtbar sind.

20. Einschränkungen und Grenzen

Im Folgenden werden die Grenzen der Studie dargestellt. Sie beziehen sich sowohl auf das konzipierte Testverfahren, als auch auf die begleitend eingesetzten Instrumente sowie die Methodik.

20.1 Das Testinstrument

Mit Blick auf die bisherige Version des Testinstrumentes ist zu kritisieren, dass das Testverfahren zu wenig Items umfasst. Da das Instrument einerseits relevante kognitive Teilkompetenzen unterschiedlicher Problemlösephasen erfassen sollte, jedoch zugleich zeitlichen Beschränkungen unterlag, musste die Itemanzahl gekürzt werden. Um eine theoretische und auch empirische Unteridentifikation zu vermeiden, empfehlen sich allerdings mehrere Indikatoren. Eine Erhöhung der Itemanzahl würde sowohl die theoretische Differenzierungsfähigkeit, als auch die empirische Berechnungsgrundlage verbessern und belastbarere Schätzungen der Zuverlässigkeit des Testverfahrens ermöglichen.

Zudem wirkt sich die geringe Itemzahl und heterogene Zusammenstellung der Items ungünstig auf die Konsistenz des Testverfahrens aus, die sich bislang lediglich als Orientierungsgröße angeben lässt. Zwar wurde das Testverfahren für die Hauptstudie nochmals überarbeitet, sodass die Auswahl der Testitems möglichst vergleichbaren und graduell abstufbaren Kriterien zur Manipulation der Schwierigkeit folgte. Bislang bietet das Testverfahren aufgrund seiner Kürze und Heterogenität jedoch lediglich einen ersten Einblick in Schwierigkeiten und Anforderungen im Zusammenhang mit fächerübergreifendem Problemlösen.

Der Blick auf die Items zum deduktiven Schließen oder zur Handlungsplanung zeigt, dass gerade hier die Heterogenität und der Schwierigkeitsgrad der Items zu Einschränkungen in der Aussagekraft führen. Geringe Itemvarianzen, die wenig Aussagen über die Streuung der Kompetenzen erlauben, verweisen darauf, dass die Testleistung teilweise sehr breit streut. Es war zwar im Sinne der Studie, auch über sehr schwierige Items im oberen Kompetenzbereich zu differenzieren; bei sehr kleinen

Variablensets wirkt sich die hohe Schwierigkeit jedoch zu Lasten der Aussagekraft aus. Um Kompetenzausprägungen verlässlich zu messen und ausreichend nicht nur zwischen Erst-, Dritt- und Sechstsemestern, sondern auch innerhalb der Semester zu diskriminieren und Boden- oder Deckeneffekte zu vermeiden, müsste die Anzahl der Items erhöht werden.

Die Befunde konnten zeigen, dass unterschiedliche Ansatzpunkte zur Variation der Schwierigkeit herangezogen werden können. Aufschlussreich wäre es dementsprechend, Variablensets zu konzipieren, die sowohl isolierte einzelne als auch kombinierte Schwierigkeitsmomente beinhalten. Eine breitere Variation des Informationsmaterials würde es erlauben, den Umgang mit unterschiedlichem Stimulusmaterial zu erfassen (Grafiken, Statistiken, textbasiertes, figuratives Material).

Einer Einschränkung unterliegt das Testverfahren auf konzeptueller Ebene auch über das Skalenniveau der Einzelitems. Die dichotome Antwortskala, die auf einer richtig-falsch Einordnung eines gegebenen Statements basiert, erlaubt zwar auf Ebene der Variablensets sowie auf Ebene der Subtests eine metrische Skalierung. Auf Einzelitemebene schränkt sie die Berechnungsmöglichkeiten jedoch ein und erhöht die Wahrscheinlichkeit einer Zufallslösung. Hintergrund für die richtig-falsch Konzeption beim Subtest Situationsanalyse waren einerseits zeitökonomische Gründe. Andererseits wurde das Antwortformat auch deshalb gewählt, um eindeutig beantwortbare Testitems (induktives, korrelatives Schließen) zu konzipieren. Partiiell abgestufte, eindeutig zu beantwortende Antwortmodalitäten stellen unter Berücksichtigung der Maßgabe der zeitlichen Ökonomie einen Ansatzpunkt zur Überarbeitung dar. Im Hinblick auf Zielevaluation und Handlungsplanung könnten die Items insofern überarbeitet werden, dass die Verknüpfung der Informationen Kompetenzen erfordert, die im Rahmen der Situationsanalyse erfasst werden. Im Zuge einer Itemsrevision kann weiterhin überlegt werden, inwieweit über die Vorgabe unterschiedlicher geschlossener Antwortmodalitäten Schwierigkeiten im logischen Schließen differenzierter gemessen werden könnten. Die Vorgabe von Mehrfachwahlantworten erhöht nicht nur die Ratewahrscheinlichkeit, sondern erlaubt auch die Überprüfung unterschiedlicher Fehlschlüsse. Die Aufdeckung von Schwierigkeiten beim Problemlösen ist letztlich die ausschlaggebende Determinante, um Problemlöseverhalten und -erfolg vorherzusagen zu können (Kotovsky & Simon, 1990).

20.2 Prüfung der konvergenten und diskriminanten Validität

Der Frage, inwieweit das Testverfahren das erfasst, was es vorgibt zu erfassen, konnte lediglich über die Augenscheinvalidität sowie über die Prüfung gemeinsamer Varianz mit Testintelligenz nachgegangen werden. Eine Prüfung diskriminanter Validität ist nur eingeschränkt über die Inbezugsetzung der Testleistung mit motivationalen Komponenten möglich, die Überprüfung der konvergenten Validität fehlt. Aussagen zur Validität sind dementsprechend nur eingeschränkt möglich. Der Vergleich mit einem belastbareren Testverfahren, das zudem unterschiedliche

Intelligenzmodalitäten umfasst, wäre aufschlussreich. Zahlreiche Befunde verweisen darauf, dass vor allem Verarbeitungskapazität eine Invariante im Zusammenhang mit Problemlösen darstellt. Die Verwendung eines breiten Intelligenzdiagnostikums wie bspw. des BIS (Jäger, 2010) würde es erlauben, Zusammenhänge und Abgrenzungen mit Invarianten des Problemlösens auch in Abhängigkeit unterschiedlichen Stimulusmaterials (figural, numerisch, symbolisch) zu untersuchen. Die Kontrolle domänenspezifischer Kompetenzen (bspw. deklaratives, prozedurales Wissen) wäre hilfreich, um das Konstrukt, bzw. Teilkomponenten fächerübergreifender Problemlösekompetenzen domänen(un)spezifisch weiter einzuordnen.

20.3 Methode und Stichprobe

Das Querschnittsdesign erlaubt zwar eine Analyse der Kompetenzausprägungen, allerdings keine Aussagen über Kausalitäten. Die Befunde und Wirkungszusammenhänge stellen lediglich Momentaufnahmen dar. Das Querschnittsdesign schränkt die Hinweise auf mögliche „Kompetenzentwicklungen“ im Laufe des Studiums ein. Längsschnittstudien oder *matching*-Verfahren würden es ermöglichen, Entwicklungsverläufe nachzuzeichnen und Kausalzusammenhänge zu beobachten. Im Rahmen von Prä-Postdesigns oder auch qualitativ ausgerichteten Settings könnte zudem differenzierteren Fragestellungen nachgegangen werden: Aufschlussreich wäre die Frage nach vorhandenen Problemlöseheuristiken oder Fehlermustern in Abhängigkeit von Disziplin und Semester: Welche Problemlösestrategien werden abgerufen? Welche Fehler treten beim logischen Schließen auf, können systematische Fehler identifiziert werden? Welche Entscheidungsregeln liegen vor? Werden sie durch Expertise beeinflusst? Qualitative Studien würden sich als methodische Herangehensweise anbieten (bspw. Introspektion). Die Orientierung am Output, das heißt die Orientierung an der richtig-falsch Beantwortung der Testitems kann keine Aufschlüsse diesbezüglich geben. Qualitative Analysen wären auch im Hinblick auf die Frage der Wirkungsweise epistemologischer Überzeugungen und der Selbstwirksamkeit aufschlussreich. Vor dem Hintergrund der Befunde stellt sich die Frage, worin sich Studierende mit reflektierteren epistemischen Überzeugungen von naiveren Studierenden unterscheiden und inwieweit die Wirkungsweise der subjektiven Überzeugungen im Zusammenhang mit Problemlösen weiter nachgezeichnet werden könnte. Prä-Postdesigns (inkl. Paralleltestverfahren) könnten konzipiert werden, um Interventionseffekte zu messen und der Frage nach dem Training fächerübergreifender Problemlösekompetenzen nachzugehen.

20.4 Kovariaten fächerübergreifenden Problemlösens

Im Rahmen der Diskussion der Befunde zu motivationalen und metakognitiven Kovariaten wurde bereits kritisch angemerkt, inwieweit die verwendeten Testinstrumente ggf. zu unspezifisch messen

(bspw. Studieninteresse). Zudem wurde dargestellt, welche weiteren Fragestellungen im Zusammenhang mit motivationalen und metakognitiven Faktoren aufschlussreich wären. An dieser Stelle wird deshalb dargestellt, inwieweit fächerübergreifende Problemlösekompetenzen über den Einbezug bislang nicht berücksichtigter Kovariaten differenzierter modelliert werden könnten.

Im Rahmen der vorliegenden Studie konnte im Hinblick auf die Frage, wie Wissen erlernt wird (Quelle des Wissens) keine direkte Einflussnahme skizziert werden. Aufschlussreich könnte es dementsprechend sein, nach deklarativem, prozeduralem und konditionalem Strategiewissen zu fragen. Bisherige Befunde verdeutlichen, dass die Überzeugung von der Sicherheit des Wissen und der Quelle des Wissens mit weniger elaborierten Lernstrategien einhergehen (u.a. Urhahne & Hopf, 2004). Die Frage wäre, inwieweit bei weniger reflektierten Studierenden überhaupt entsprechendes Wissen und die Überzeugung vorhanden ist, aktiv in den Konstruktionsprozess der Wissensentwicklung eingebunden zu sein. Der Einfluss der Selbstwirksamkeit legt nahe, Fragen danach nachzugehen, inwieweit Wissen um Strategien des Wissenserwerbs vorhanden sind und inwieweit Studierende der Überzeugung sind, dass sie eigentätig am Prozess der Wissensentwicklung teilnehmen können, das heißt, die Strategien des Wissenserwerbs auch einsetzen. Diese Frage würde auch Aufschluss darüber bieten, inwieweit die Universität gegebenenfalls eher zu mechanischem Lernen animiert und tiefenorientierte Lernstrategien nicht unbedingt mit besseren Lernergebnissen und Noten zusammenhängen, allerdings mit besseren Ergebnissen im Problemlösen – wie es Benaware und Deci (1984) vermuten. Daneben bleibt auch die Frage der Vermittlung von Transferkompetenzen weiter zu klären. Die Studie konnte zeigen, dass die Überzeugung, dass domänenspezifische Problemlösemethoden auch auf fächerübergreifende Probleme angewendet werden können, signifikant über die Semester hinweg abnimmt und im Weiteren zwar bedeutsam mit Selbstwirksamkeit, allerdings nicht mit fächerübergreifendem Problemlösen zusammenhängt. Die Frage ist, ob der Einfluss auf fächerübergreifendes Problemlösen ggf. über die Überzeugung moderiert wird, dass Studierende mit Wissen eigentätig umgehen können.

Schließlich bleibt zu klären, welche Rolle die Überzeugungen der Lehrenden spielen. Pädagogisches Wissen ist ein relevanter Einflussfaktor nicht nur auf Lernen, sondern auch auf die Fähigkeit, Probleme zu lösen (Carpenter, Fennema, Peterson & Carey, 1988; für einen Überblick auch Hofer 2008; Calderheard, 1996; Pajares, 1996; Shavelson & Stern, 1981). Carpenter et.al. (1988) konnten auch beobachten, dass Lehrende nicht wussten, welche Problemlösestrategien von ihren Schülern verwendet werden. Die Autoren betonen, dass ihre Befunde sich in die zahlreichen Ergebnisse einordnen lassen, die bekräftigen, dass Lehrende eine ungenügende Wissensbasis besitzen, um die Problemlöseprozesse ihrer Schüler zu diagnostizieren und auf dieser Basis ihre Lehre zu konzipieren. Es stellt sich dementsprechend die Frage, inwieweit das Zusammenspiel zwischen subjektiven Überzeugungen der Lehrenden und Studierenden sowie das Wissen um Problemlösestrategien und das Wissen um Problemlösekompetenzen in seiner Wechselwirkung weiter nachgezeichnet werden kann.

20.5 Grenzen und Einschränkungen im Hinblick auf Schlussfolgerungen für die Lehre

Im Hinblick auf die Lehre wurden Ansatzpunkte für die Konzeption von Lehrsettings bereits beschrieben. Einschränkungen unterliegt das Testverfahren in der Hinsicht, dass psychometrisch konzipierte Testaufgaben für die Entwicklung von Unterrichtsqualität relativ uninteressant sind, da ihre Qualität an anderen Kriterien gemessen wird als die Qualität von Lernaufgaben (Schecker & Parchmann, 2006). Da Testverfahren als diagnostische Instrumente allerdings für die Gestaltung von Lehr-Lernprozessen genutzt werden können, ist mit Blick auf das vorliegende Testverfahren kritisch anzumerken, dass die Verknüpfung zwischen Testitems und Lernaufgaben bei einer Überarbeitung spezifischer mitgedacht werden könnte.

21. Zusammenfassung und Ausblick

Ausgangspunkt der Studie war, dass das Thema der fächerübergreifenden Problemlösekompetenzen einen zentralen Platz in der bildungspolitischen Debatte einnimmt. Im beruflichen Alltag müssen Studierende mit Problemstellungen umgehen können, die neuartig und komplex sind und den Transfer domänenspezifischer Kompetenzen erfordern. Eine Erwerbstätigenbefragung des Bundesinstitutes für Bildung und Beruf sowie der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin (BAuA, 2006) zeigt, dass mehr als 75% des beruflichen Alltags damit verbracht wird, auf unvorhergesehene Probleme zu reagieren, schwierige Entscheidungen zu treffen oder auch schwierige Sachverhalte zu vermitteln (auch Einleitung, Kap. 1). Viele Studierende im Bachelorbereich fühlen sich wenig vorbereitet auf den Umgang mit diesen Anforderungen (Briedis, Heine, Konegen-Grenier & Schröder, 2011; Multrus, 2011; Ramm & Simeaner, 2010). Die Frage nach fächerübergreifenden Problemlösekompetenzen ist jedoch nicht nur relevant, da Problemlösen als Bildungsergebnis wichtig ist. Vielmehr sind fächerübergreifende Problemlösekompetenzen auch bedeutsam, da Problemlösen eine Voraussetzung für erfolgreiches weiteres Lernen darstellt (Fleischer et. al. 2010; Baumert et. al., 2001).

Die zentrale Frage der Studie lautete dementsprechend, wie fächerübergreifende Problemlösekompetenzen theoretisch beschrieben und empirisch überprüft werden können, welche Kompetenzen bei Studierenden unterschiedlicher Fächer und Semester beim Lösen neuartiger Probleme vorhanden sind und in welchem Wirkungs- und Bedingungs Zusammenhang sie mit Faktoren wie metakognitiven, subjektiven Überzeugungen und dem am Studium gezeigten Interesse stehen. Die für den Lehralltag daraus abgeleitete Frage war, wie Studierende beim Lösen unbekannter und neuer Probleme unterstützt werden können, um vor dem Hintergrund der Bedeutsamkeit fächerübergreifender Problemlösekompetenzen dem bisherigen subjektiven Erleben der Studierenden entgegenzuwirken und sowohl Problemlösefähigkeiten als Bildungsergebnis, als auch als Voraussetzung für weiteres Lernen zu fördern. Die Studie sollte insgesamt Ansatzpunkte zur Identifikation von *academic skills* bieten.

Bei der Auseinandersetzung mit diesen Fragen waren folgende Forschungslücken zu entdecken:

Erstens, das Kompetenzkonstrukt, insbesondere das Konstrukt des fächerübergreifenden Problemlösens ist theoretisch und empirisch unscharf (u.a. Leuders & Sodian, 2013). Zweitens, es gibt keine validen Messinstrumente zur Erfassung fächerübergreifender Kompetenzen (u.a. Blömeke et. al., 2013; Zlatkin-Troitschanskaia & Kuhn, 2010). Als Folge dessen liegt drittens bislang auch keine Beschreibung von Stärken und Schwächen Studierender beim fächerübergreifenden Problemlösen vor. Nach der eingehenden Beschäftigung mit fächerübergreifendem Problemlösen im Rahmen dieser Arbeit kann im Hinblick auf die theoretische und empirische Unschärfe des Konstruktes festgehalten werden, dass Problemlösekompetenzen mithilfe einer Orientierung an Befunden zum logischen Schließen und informationsverarbeitungstheoretisch begründeten Befunden zu Problemlöseprozessen unter Berücksichtigung von Problemcharakteristika gut operationalisiert werden konnten. Der Aspekt des Fächerübergreifenden wurde mit einer Orientierung an Befunden des analogen Problemlösens und der Expertise- und Transferforschung fassbar. In der Folge konnte ein pragmatisches und funktionales Modell erarbeitet werden, das eine Messbarkeit und Systematisierung des Konstruktes gewährleistete und letztlich über das Studium vermittelte Stärken der Studierenden sowie bis zum sechsten Semester vorhandene Schwächen in Abhängigkeit der Konzeptualisierung des Problems identifizieren konnte.

Mit Blick auf die einleitend dargestellte Diskussion um *academic skills* und die Befunde, dass viele Studierende, insbesondere Bachelorstudierende der Überzeugung sind, wenig auf ihren künftigen Beruf und fächerübergreifendes Problemlösen vorbereitet zu werden (Grützmaker, Ortenburger & Heine, 2011; Multrus, 2011; Simeaner, Ramm & Kolbert-Ramm, 2010; Allen & van der Velden, 2007; Teichler & Schomburg, 2006; Borden & Rajackie, 2000), ist abschließend noch Folgendes festzuhalten: Erstaunlich ist, dass die Studierenden im Laufe ihres Studiums im Lösen bestimmter fächerübergreifender Probleme bedeutsam besser und in Bezug auf die Einschätzung von Wissen und der Natur des Wissens reflektierter werden. Gleichzeitig sind sie aber immer weniger der Überzeugung, dass sie unter Anwendung der fachspezifisch erworbenen Problemlösekompetenzen fächerübergreifend (oder außerhalb ihrer Disziplin) Probleme lösen können. Daraus lässt sich die Frage ableiten, inwieweit Studierende der Überzeugung sind, in ihrem Studium ausschließlich zu Fachspezialisten zu werden? Haben sie den Eindruck, sich zu lange mit ihrer Disziplin befasst zu haben – tatsächlich werden sie aber in der Anwendbarkeit und Übertragbarkeit ihrer Kompetenzen durchaus besser? Für Studium und Lehre lässt sich daraus ableiten, dass dieses Missverhältnis angesprochen werden sollte. Es könnte eine Erklärung dafür sein, warum Studierende äußern, sie hätten im Studium wenig für ihren Beruf gelernt. Das heißt, diese ggf. vorhandenen kognitiven Dissonanzen sollten im Studienverlauf thematisiert werden, Studierende sollten dazu angehalten werden, ihren Eindruck zu überprüfen und ihre fächerübergreifenden und metakognitiven Kompetenzen zu evaluieren. Dies könnte verschiedene positive Konsequenzen haben: Erstens eine angemessenere Bewertung des eigenen Studiums im Rückblick, zweitens mehr Selbstsicherheit beim

Eintritt in den Beruf, drittens ggf. einen Einfluss auf das Studieninteresse, das in der vorliegenden Konzeption in den höheren Semestern abnimmt. Eine Beschäftigung mit der einleitend dargestellten Frage, was ein Studium ausmacht und welche *academic skills* von den Studierenden erworben werden sollen, erscheint dementsprechend nicht nur für Politik, Lehrende und Arbeitgeber von Bedeutung, sondern insbesondere auch für die Studierenden.

Literaturverzeichnis

Abbott, I. (1997). Why do we have to do key skills? Student views about General National Vocational Qualification. *Journal of Vocational Education and Training*, 49, 4, 617-630.

Adler, J. E. & Rips, L. J. (2008). *Reasoning: Studies of Human Inference and Its Foundation*. Cambridge, UK: Cambridge University Press.

Albanese, M. A. & Mitchell, S. (1993). Problem-based learning: A review of literature of its outcome and implementation issue. *Academic Medicine*, 68, 52-81.

Alexander, P. A. (2003). Can we get there from here? *Educational Researcher*, 32, 3-4.

Allen J., Ramaekers, G. & van der Velden, R. (2005). Measuring Competencies of Higher Education Graduates. In D. J. Weerts & J. Vidal (Hrsg.), *Enhancing Alumni Research: European and American Perspectives, New Directions for Institutional Research* (49-59). San Francisco: Jossey-Bass.

Amelang, M. & Funke, J. (2005). Entwicklung und Implementierung eines kombinierten Beratungs- und Auswahlverfahrens für die wichtigsten Studiengänge an der Universität Heidelberg. *Psychologische Rundschau*, 56, 135-137.

American Educational Research Association, American Psychological Association & National Council for Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, D.C.: American Educational Research Association.

Anand, P. G. & Ross, M. (1987). Using computer-assisted instruction to personalize arithmetic materials for elementary school children. *Journal of Educational Psychology*, 79, 72-78.

Anderson, J. R. (1993). Problem Solving and Learning. *American Psychologist*, 48, 1, 35-44.

Anderson, J. R., Reder, L. M. & Simon, H.A. (1996). Situated learning and education. *Educational Researcher*, 25, 4, 5-96.

Anderson, J. R. (2007). *Kognitive Psychologie*. Heidelberg: Springer.

- Andre, D. & Fernand, G. (2008). Sherlock Holmes – an experts view of expertise. *British Journal of Psychology*, 99, 109-125.
- Arbuckle, J. L. & Wothke, W. (1999). *AMOS 4.0 User's Guide*. Chicago: Small Water.
- Arnold, R. (2002). Von der Bildung zur Kompetenzentwicklung. In E. Nuissl, Ch. Schiersmann & H. Siebert (Hrsg.), *Literatur- und Forschungsreport Weiterbildung* (26-38). Bielefeld: Bertelsmann.
- Artelt, C., Stanat, P., Schneider, W. & Schiefele, U. (2001). Lesekompetenz: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (69-137). Opladen: Leske + Budrich.
- Artelt, C., Schiefele, U. & Schneider, W. (2001). Predictors of reading literacy. *European Journal of Psychology of Education*, 16, 363-383.
- Atkinson, R. K., Derry, S. J., Renkl, A. & Wortham, D. (2000). Learning from Examples: Instructional Principles from the Worked Examples Research, *Review of Educational Research*, 70, 2, 181-214.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2006). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung*. Springer.
- Baddeley, A.D. (1986). *Working memory*. Oxford UK: Oxford University Press.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillman, K.-J., & Weiß, M. (2001). PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. Opladen: Leske + Budrich.
- Baumert, J., Bos, W., Brockmann, J., Gruehn, , Klieme, E., Köller, O., Lehmann, R., Lehrke, M., Neubrand, J., Schnabel, K. U. & Watermann, R. (2000). *TIMSS III Deutschland. Der Abschlussbericht. Zusammenfassung ausgewählter Ergebnisse der Dritten Internationalen Mathematik- und Naturwissenschaftsstudie zur mathematischen und naturwissenschaftlichen Bildung am Ende der Schullaufbahn*. Berlin.

Barnett, R. (1994). *The limits of competence: Knowledge, Higher Education, and Society*. Open University Press and Society for Research into Higher Education.

Barnett, R. (2000a). *Realizing the University in an age of supercomplexity*. Ballmoor, Philadelphia: SRHE, Open University Press.

Barnett, R. (2000b). Supercomplexity and the Curriculum. *Studies in Higher Education*, 25, 3, 255-265.

Barnett, R. (2000c). University knowledge in an age of supercomplexity. *Higher Education*, 40, 409-422.

Barnett, M. & Ceci, J. (2002). When and Where Do We Apply What We Learn? A Taxonomy for Far Transfer. *Psychological Bulletin*, 128, 4, 612–637.

Barron, J., Schwartz, D. L., Vye, N. L., Moore, A., Petrosino, A., Zech, L., Bransford, J. D. & The Cognition and Technology Group at Vanderbilt. (1998). Doing with understanding: Lessons from research on problem- and project-based learning. *Journal of the Learning Sciences*, 7, 3-4, 271-311.

Barrows, H. (2000). *Problem-Based Learning Applied to Medical Education*, Southern Illinois University Press, Springfield.

Bartosch, U. Maile, A. & Speth, Ch. (2008). Qualifikationsrahmen Soziale Arbeit (QR SArb). Zugriff am 12.01.2014 unter http://www.fbts.de/fileadmin/fbts/Aktuelles/QRSArb_Version_5.1.pdf

Bassok, M. (2003). Analogical Transfer in Problem Solving. In J. E. Davidson & R. J. Sternberg, *The Psychology of Problem Solving* (343-369). Cambridge: University Press.

Baxter Magolda, M. B. (1992). *Knowing and reasoning in college: Gender related patterns in students' intellectual development*. San Francisco, CA: Jossey-Bass.

Beauducel, A. & Wittmann, W. W. (2005). Simulation study on fit indices in confirmatory factor analysis based on data with slightly distorted simple structure. *Structural Equation Modeling*, 12, 41-75.

- Bearman, Ch., Mall, L. J. & Ormerod, Th. C. (2007). The structure and function of spontaneous analogizing in domain-based problem solving. *Thinking and reasoning*, 13, 3, 272-294.
- Beck, B. & Klieme, E. (2007). *Sprachliche Kompetenzen. Konzepte und Messung. DESI-Studie (Deutsch Englisch Schülerleistung International)*. Weinheim: Beltz.
- Bendixen, L. D. & Rule, D. C. (2004). An Integrative Approach to Personal Epistemology: A Guiding Model, *Educational Psychologist*, 39, 1, 69-80.
- Bennet, N., Dunne, E. & Carré, C. (1999). Patterns of core and generic skill provision in higher education. *Higher Education*, 37, 71-93.
- Bennet, N., Dunne, E. & Carré, C. (2000). *Skills development in Higher Education and Employment*. Ballmoor, Philadelphia: SRHE, Open University Press.
- Bennett, R. E., Perskey, H., Weiss, A. & Jenkins, F. (2010). Measuring Problem Solving with Technology: A Demonstration Study for NAEP, *The Journals of Technology, Learning, and Assessment*, 8, 8. Zugriff am 12.01.2014 unter <http://www.jtla.org>
- Belenky, M., Clinchy, B., Goldberger, N. & Tarule, J. (1986). *Women's Ways of Knowing*. New York: Basic Book.
- Belland, B. R., French, B. F. & Ertmer, P. A. (2009). Validity and Problem-Based Learning Research: A Review of Instruments Used to Assess Intended Learning Outcomes. *Interdisciplinary Journal of Problem-based Learning*, 3, 1, 59-89.
- Beller, S., Kuhnmünch, G., Spada, H. (o. J.). Logisches Schließen ohne 'Biases'. Zugriff am 12.01.2014 unter <https://www.psychologie.uni-freiburg.de/Members/beller/publications1/sb01teap>
- Berdahl, R. (o. J.). Thoughts About Academic Freedom, Autonomy and Accountability. Zugriff am 12.01.2014 unter http://www.magna-charta.org/library/userfiles/file/Berdahl_2010-Thoughts_Abou_Academic_Freedom,_Autonomy_and_Accountability.pdf
- Berliner, D. C. & Calfee, R. C. (1996). *Handbook of educational Psychology*. New York: Simon & Schuster McMillan.

- Biggs, J. (2003). *Teaching for quality learning at university*. Buckingham: The Society for Research into Higher Education & Open University Press.
- Blanchette, I. & Leese, J. (2010). The Effect of Negative Emotion on Deductive Reasoning. *Experimental Psychology*, 58, 3, 235-246.
- Blanchette, I. & Leese, J. (2011). Physiological arousal and logicity: The effect of emotion on conditional reasoning. *Experimental Psychology*, 58(3), 235–246.
- Blumberg, P. & Michael, J. A. (1991). The development of selfdirected learning behavior in a partially teacher-centered problem-based learning curriculum. *Teaching and Learning Medicine*, 4, 3-8.
- Blömeke, S. & Zlatkin-Troitschanskaia, O. (2013). *Kompetenzmodellierung und Kompetenzerfassung im Hochschulsektor: Ziele, theoretischer Rahmen, Design und Herausforderungen des BMBF-Forschungsprogramms KoKoHs (KoKoHs Working Papers, 1)*. Berlin & Mainz: Humboldt-Universität & Johannes Gutenberg-Universität.
- Blömeke, S., Zlatkin-Troitschanskaia, O., Kuhn, C. & Fege, J. (2013). *Modeling and Measuring Competencies in Higher Education*. Rotterdam: Sense Publisher.
- Bransford, J. D. & Stein, B. (1993): *The IDEAL problem solver*. New York, Freeman.
- Bransford, J. D., Brown, A. L. & Cocking, R. R. (2000). *How People learn. Brain, Mind, Experience and School*. Washington: National Academy Press.
- Bransford, J. D. & Johnson, M. K. (1972). Contextual Prerequisites for Understanding: Some Investigations of Comprehension and Recall. *Journal of verbal Learning and verbal Behavior*, 11, 717-726.
- Briedis, K., Heine, Ch., Konege-Grenier, Ch. & Schröder, A.-K. (2011). *Mit dem Bachelor in den Beruf. Arbeitsmarktbefähigung und -akzeptanz von Bachelorstudierenden und -absolventen*. Stifterverband für die Deutsche Wissenschaft. Essen.
- Brimm, O. G., Glass, D.C., Lavin, D.E. & Goodman, N. (1962). *Personality and decision processes: Studies in the social psychology of thinking*. Stanford, CA: Stanford University Press.

- Bridges, D. (2000). Back to the Future: the higher education curriculum in the 21st century. *Cambridge Journal of Education*, 30, 1, 37-55.
- Britton, B. K. & Graesser, A. C. (1996). *Models of understanding text*. Mahwah, NJ: Erlbaum.
- Brödel, R. (2002). Relationierungen zur Kompetenzdebatte. In E. Nuissl, Ch. Schiersmann & H. Siebert (Hrsg.). *Literatur- und Forschungsreport Weiterbildung* (39-47). Bielefeld: Bertelsmann.
- Brophy, D. R. (1998). Understanding, Measuring, and Enhancing Individual Creative Problem-Solving Effort. *Creativity Research Journal*, 11, 2, 123-150.
- Brown, D. E. & Clement, J. (1989). Overcoming misconceptions via analogical reasoning: abstract transfer versus explanatory model construction. *Instructional Science*, 18, 237-261.
- Brown, J., Collins, A. & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18, 1, 32-42.
- Brunstein, A. (2001). *Zum Einfluss der visuellen Oberfläche auf bildhafte Analogien*. Dissertation, Universität Hamburg. Zugriff am 19.01.2014 unter <http://ediss.sub.uni-hamburg.de/volltexte/2001/510/>
- Buehl, M. (2008). Assessing the multidimensionality of students epistemic beliefs across diverse culture. In M. Khine (Hrsg.), *Knowing, knowledge and beliefs: Epistemological studies across diverse cultures* (65-112). Netherlands: Springer.
- Bulheller, S. & Häcker, H. (2005). *Advanced Progressive Matrices (APM) - Manual*. Frankfurt/Main: Swets Test Service.
- Bundesvereinigung der deutschen Arbeitgeberverbände (2008). *Bildung schafft Zukunft. Beschäftigungsfähigkeit von Hochschulabsolventen stärken. Für eine bessere Arbeitsmarktrelevanz des Hochschulstudium*. Gemeinsames Memorandum von BDA, BDI und HRK.
- Bundesministerium für Bildung und Forschung, Bundesvereinigung der deutschen Arbeitgeberverbände & Bundesverband der deutschen Industrie (2007). *Fit für den Job?! Arbeitsmarktcompetenzen der Studierenden stärken*. Berlin.

- Bundesamt für Statistik (2008). *Schlüsselkompetenzen der Schweizer Hochschulabsolvent/innen. Thematischer Sammelband mit empirischen Ergebnissen der Absolventenstudie*. Neuchâtel.
- Buehner, M. J. & Cheng, P. W. (2005). Causal learning. In K. J. Holyoak, R. G. Morrison (Hrsg.), *The Cambridge Handbook of Thinking and Reasoning* (143-168), Cambridge: University Press.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson.
- Butterfield, E. C. & Nelson, G. D. (1989). Theory and practice of teaching for transfer. *Educational Technology Research and Development*, 37, 3, 5-38.
- Bollen, K. A. (1989). *Structural Equations with Latent Variable*. New York: Wiley.
- Bos, W., Lankes, E. M., Prenzel, M., Schwippert, K., Valtin, R. & Walther, G. (2008) *IGLU. Vertiefende Analysen zu Leseverständnis, Rahmenbedingungen und Zusatzstudien*. München: Waxmann.
- Bortz J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human und Sozialwissenschaftler*. Heidelberg: Springer.
- Bransford, J. D., Brown, A. L. & Cocking, R. R. (2000). *How People Learn. Brain, Mind, Experience and School*. Washington: National Academy Press.
- Bransford, J., Vye, N., Stevens, R., Kuhl, P., Schwartz, D., Bell, Ph., Meltzoff, A., Barron, B., Pea, R., Reeves, B., Roschelle J. & Nora Sabelli (2006). Learning Theories and Education: Toward a Decade of Synergy. In P. Alexander & P. Winne (Hrsg.), *Handbook of Educational Psychology* (209-244). Mahwah, NJ: Erlbaum.
- Briedis, K. (2007): Die HIS Absolventenstudie. In Hochschulrektorenkonferenz (Hrsg.), *Potentiale von Absolventenstudien für die Hochschulentwicklung, Dokumentation einer Veranstaltung der HRK in Kooperation mit dem INCHER-Kassel und dem Arbeitsbereich Absolventenforschung der FU Berlin am 18. und 19. Mai 2006 an der Universität Kassel, Bonn, März 2007*. Beiträge zur Hochschulpolitik.
- Bromley, D. B. (1988). *Human ageing*. Harmondsworth: Penguin.

- Bromme, R., Pieschel S. & Stahl, E. (2010). Epistemological beliefs are standards for adaptive learning: a functional theory about epistemological beliefs and metacognition. *Metacognition Learning*, 5, 7-26.
- Bruner, J., Goodnow, J. J. & Austin, G. A. (1956). *A study of thinking*. New York: Wiley.
- Byrne, B. M. (2001). Structural Equation Modeling with AMOS, EQS, and LISREL: Comparative Approaches to Testing for the Factorial Validity of a Measuring Instrument,” *International Journal of Testing*, 1, 1, 55-86.
- Calderhead, J. (1996). Teachers: Beliefs and Knowledge. In D. C. Berliner & R. C. Calfee (Hrsg.), *Handbook of educational Psychology* (709-725). New York: Simon & Schuster McMillan.
- Carey, S. & Spelke, E. (2008). Domain-specific knowledge and conceptual change. In J. E. Adler & L. J. Rips (Hrsg.), *Reasoning: Studies of Human Inference and Its Foundation* (807-826). Cambridge, UK: Cambridge University Press.
- Caprara, G. V., Vecchione, M., Alessandri, G., Berbinò, M. & Barbaranelli, C. (2011). The contribution of personality traits and self-efficacy beliefs to academic achievement: A longitudinal study. *British Journal of Educational Psychology*, 81, 78-96.
- Chapman, L. J. & Chapman, J. P. (1959). Atmosphere effect reexamined. *Journal of Experimental Psychology*, 58, 220-226.
- Catrambone, R. (1996). Generalizing solution procedures learned from example *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 4, 1020-1031.
- Catrambone, R. (o. J.). Teaching subgoals to students improves learning. Zugriff am 12.01.2014 unter http://www.iwm-kmrc.de/workshops/SIM2004/pdf_files/Catrambone.pdf
- Carpenter, T. P., Fennema, E., Peterson, P. L. & Carey, D. A. (1988). Teachers' pedagogical content knowledge of students' problem solving in elementary arithmetic. *Journal for Research in Mathematics Education*, 19, 5, 385-401.
- Caplan, J. B. & Caplan, P. J. (2005). The persuasive search for sex differences in mathematical ability In A. M. Gallagher & J. C. Kaufman (Hrsg.), *Gender Differences in Mathematics: An integrative psychological approach*. Cambridge: Cambridge University Press.

Center for educational research and innovation (1995). *Measuring what students learn*. Organisation for Economic Co-Operation and Development.

Chai, Ch. & Kine, M. (2008). Assessing the Epistemological and Pedagogical Beliefs among Pre-Service Teachers in Singapore. In M. Khine (Hrsg.), *Knowing, Knowledge, and Beliefs: Epistemological Studies across Diverse Cultures* (287-299). Springer.

Chai, Ch., Khine, M. & Teo, T. (2006). Epistemological beliefs on teaching and learning: A survey among pre-service teachers in Singapore. *Educational Media International*, 43, 285-298.

Chase, W. G. & Simon, H. A. (1973) Perception in chess. *Cognitive Psychology*, 4, 55-81.

Cheng, P. & Holyoak, K. J. (1985). Pragmatic reasoning schemas, *Cognitive Psychology*, 17, 391-416.

Cheng, P. & Holyoak, K. J. (2008). Pragmatic reasoning Schemas. In J. E. Adler & L. J. Rips (Hrsg.), *Reasoning: Studies of Human Inference and Its Foundation* (827-842). Cambridge, UK: Cambridge University Press.

Chi, M. T., Glaser, R. & Rees, E. (1982). Expertise in Problem solving. In R. J. Sternberg (Hrsg.), *Advances in the psychology of human intelligence* (7-76). Hillsdale, NJ: Erlbaum.

Chi, M. T. H., Feltovich, P. & Glaser, R. (1981). Categorization and representation of physics problems by experts and novice. *Cognitive Science*, 5, 121–152.

Chi, M. T., Glaser, R. & Farr, M. J. (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.

Chi, M. T. H. (2006). Laboratory Methods for Assessing Experts' and Novices' Knowledge. In K.A. Ericsson, N. Charness, P. Feltovich & R. Hoffman (Hrsg.), *Cambridge Handbook of Expertise and Expert Performance* (121-30). Cambridge University Press.

Clanchy, J. & Ballard, B. (1995). Generic Skills in the Context of Higher Education. *Higher Education Research and Development*, 14, 2, 155-166.

Clement, C. A. & Falmange, R. J. (1986). Logical reasoning, world knowledge, and mental imagery: Interconnections in cognitive processes. *Memory & Cognition*, 14, 4, 299-307.

- Cofer, Ch. N. (1967). Conditions for the use of verbal association. *Psychological Bulletin*, 68, 1, 1-12.
- Cognition and Technology Group at Vanderbilt (1990). Anchored instruction and its relationship to situated cognition. *Educational Researcher*, 19, 6, 2-10.
- Cognition and Technology Group at Vanderbilt. (1992). The Jasper series as an example of anchored instruction: Theory, program description, and assessment data. *Educational Psychologist*, 27, 3, 291-315.
- Cohen, N., Henik, A. & Mor, N. (2011). Can Emotion modulate Attention? Evidence for Reciprocal Links in the Attentional Network test. *Experimental Psychology*, 58, 3, 171-179.
- Connell, M. W., Sheridan, K. & Gardner, H. (2003). On Abilities and Domain. In R. J. Sternberg & E. L. Grigorenko (Hrsg.), *The psychology of abilities, competencies, and expertise* (126-155). Cambridge University Press.
- Collins, A., Brown, J., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator*, 15, 6-11, 38-46.
- Csapó, B. (2004). Knowledge and competencies. In J. Letschert (Hrsg.), *The integrated person. How curriculum development relates to new competencies* (35-49). Enschede: CIDREE.
- Cummins, D. (1995). Naive theories and causal deduction. *Memory & Cognition*, 23, 646-658.
- Davidson, J. E. & Sternberg, R. J. (2003). *The Psychology of Problem Solving*. Cambridge: University Press.
- Davidson, J. E. (2003). Insights about insightful Problem Solving. In J. E. Davidson & R. J. Sternberg (2003). *The Psychology of Problem Solving* (176-206). Cambridge: University Press.
- De la Harpe, B., Radloff, A. & Wyber, J. (2000). Quality and Generic Professional Skill. *Quality in Higher Education*, 6, 3, 231-243.

- De Young, C. G., Flanders, J. L. & Peterson, J. B. (2008). Cognitive Abilities in Insight Problem Solving. An Individual Differences Problem. *Creativity Research Journal*, 20, 3, 278-290.
- De Grave, W., Boshuizen, H. P. A. & Schmidt, H. G. (1996). Problem based learning: Cognitive and metacognitive processes during problem analysis. *Instructional Science*, 24, 321-341.
- De Groot, A. (1965). *Thought and choice in chess*. The Hague: Mouton.
- Dewey, J. (1933): *How we think*. Boston, MA: D.C. Heath.
- Diamantopoulos, A. & Siguaw, J. (2000). *Introducing LISREL*. London: SAGE.
- Diehl, J. M. & Arbinger, R. (1990). *Einführung in die Inferenzstatistik*. Eschborn b. Frankfurt: Klotz.
- Diehl, J. M. & Staufenbiel, Th. (2007). *Statistik mit SPSS Version 15*. Magdeburg: Klotz.
- Dods, R. F. (1996). An action research study of the effectiveness of problem-based learning in promoting the acquisition and retention of knowledge. *Journal for the Education of the Gifted*, 19, 3, 423-437.
- Dominowski, R. L. & Bourne, L. E. Jr. (1994). History of Research on Thinking and Problem Solving. In R. J. Sternberg (Hrsg.), *Thinking and Problem Solving* (1-36). London, Kalifornien: Academic Press.
- Domsch, M. (1990). Fallstudien. In W. Sarges (Hrsg.), *Management-Diagnostik* (484-490). Göttingen, Hogrefe.
- Dörner, D. (1976). *Problemlösen als Informationsverarbeitung*. Stuttgart: Kohlhammer.
- Dörner, D. (1989). *Die Logik des Misslingens. Strategisches Denken in komplexen Situationen*. Hamburg: Rowohlt.
- Dossey, J., Csapo, B., De Jong, T., Klieme, E. & Vosniadou, St. (o. J.). *Cross-Curricular Competencies in PISA. Towards a Framework for Assessing Problem solving skill*. The INES Compendium.
- Dossey, J., Hartig, J., Klieme, E. & Wu, M. (2004). *Problem solving for tomorrow's world. First measures of cross-curricular competencies from PISA 2003*. Paris: OECD.

Downing, M. & Haladyna, T. M (2006). *Handbook of test development*. New Jersey: Lawrence Erlbaum.

Drummond, I., Nixon, I. & Wiltshire, J. (1998). Personal transferable skills in higher education: the problems of implementing good practice. *Quality Assurance in Higher Education*, 6, 1, 19-27.

Duden (2013). Kompetenz. Zugriff am 12.01.2014 unter <http://www.duden.de/rechtschreibung/Kompetenz>

Duit, R., Häußler, P. & Prenzel, M. (2001). Schulleistungen im Bereich der naturwissenschaftlichen Bildung. In F. Weinert (Hrsg.), *Leistungsmessung in Schulen* (169-186). Weinheim: Beltz.

Duncker, K. & Lees, L. S. (1945) On problem-solving. *Psychological Monographs*, 58, 5, 1-113.

Ekman, R. & Pelletier, St. (2008). Assessing student learning. A work in Progress. *Change*, 40, 4, 14-19.

Ericsson, K. A. & Smith J. (1991). *Toward a general theory of expertise: Prospects and limits*. Cambridge: Cambridge University Press.

Ericsson, K. A. & Hastie, R. (1994). Contemporary Approaches to the Study of Thinking and Problem Solving. In R. J. Sternberg (Hrsg.), *Thinking and Problem Solving* (37-82). London, Kalifornien: Academic Press.

Ericsson, K. A. (2003a). The Search for general Abilities and Basis Capacities: Theoretical Implications from Modifiability and Complexity of Mechanisms Mediating Expert Performance. In R. J. Sternberg & E. L. Grigorenko (Hrsg.), *The psychology of abilities, competencies, and expertise* (93-125). Cambridge University Press.

Ericsson, K. A. (2003b). The Acquisition of Expert Performance as Problem Solving: Construction and Modification of Mediating Mechanisms of deliberate Praxis. In Davidson & R. J. Sternberg (Hrsg.), *The Psychology of Problem Solving* (31-86). Cambridge: University Press.

- Ericsson, K. A. & Smith, J. (1991). Prospects and limits of empirical study of expertise: An introduction. In K. A. Ericsson & J. Smith (Hrsg.), *Towards a general theory of expertise: Prospects and limits* (1-38). Cambridge: Cambridge University Press.
- Ericsson, K. A. & Crutcher, R. J. (1990). The nature of exceptional performance. In P. B. Baltes, D. L. Featherman & R. M. Lerner (Hrsg.), *Lifespan development and behavior* (187-219). Hillsdale NJ: Erlbaum.
- Ericsson, K. A. & Simon, H. A. (1980). Verbal Reports as Data. *Psychological Review*, 87, 3, 215-251.
- Ernst, G. W. & Newell, A. (1969). *GPS: A case study in generality and problem solving*. New York: Academic Press.
- Erpenbeck, J. & Rosenstiel, L. v. (2007a). Einführung. In J. Erpenbeck & L. v. Rosenstiel (Hrsg.), *Handbuch Kompetenzmessung*. Stuttgart: Schäffer-Poeschel.
- Erpenbeck, J. & Rosenstiel, L. v. (2007b). *Handbuch Kompetenzmessung*. Stuttgart: Schäffer-Poeschel.
- Evans, J. St. B. T., Newstead, E. & Byrne, R. M. J. (1993). *Human reasoning*. Hove, UK: Lawrence Erlbaum.
- Evans, J. St. B. T. (2005). Deductive Reasoning. In K. J. Holyoak & R. G. Morrison (Hrsg.), *The Cambridge Handbook of Thinking and Reasoning* (169-184), Cambridge: University Press.
- Falmagne, R. J. & Gonsalves, J. (1995). Deductive Inference. *Annual Review of Psychology*, 46, 525-559.
- Fennekel, G. P. & D'Souza, (1999). *Management-Fallstudien (MFA). Handanweisung*. Göttingen: Hogrefe.
- Fleischer, J., Wirth, J., Rumann & Leutner, D. (2010). Strukturen fächerübergreifender und fachlicher Problemlösekompetenz – Analyse von Aufgabenprofilen. In E. Klieme, D. Leutner & M. Kenk (Hrsg.), *Kompetenzmodellierung Eine aktuelle Zwischenbilanz des DFG-Schwerpunktprogramms* (56. Beiheft der Zeitschrift für Pädagogik, 239–248). Weinheim: Beltz.

Fleischer, J., Koeppen, K., Martina Kenk, M., Klieme, E. & Leutner, D. (2013). Kompetenzmodellierung: Struktur, Konzepte und Forschungszugänge des DFG Schwerpunktprogramms. *Zeitschrift für Erziehungswissenschaft*, 16, 5-22.

Funke (2003). *Problemlösen*. Stuttgart: Kohlhammer.

Funke, J. & Spering, M. (2006). Methoden der Denk- und Problemlöseforschung. In J. Funke (Hrsg.), *Denken und Problemlösen* (=Enzyklopädie der Psychologie, Themenbereich C: Theorie und Forschung, Serie II: Kognition, Band 8: Denken und Problemlösen, 647-744). Göttingen: Hogrefe.

Funke, J. (2006). *Denken und Problemlösen* (=Enzyklopädie der Psychologie, Themenbereich C: Theorie und Forschung, Serie II: Kognition, Band 8). Göttingen: Hogrefe.

Funke, J. & Frensch, Peter. A. (2006). *Handbuch der Psychologie*. Göttingen: Hogrefe.

Funke, J. (2000). Psychologie der Kreativität. In R. M. Holm-Hadulla (Hrsg.), *Kreativität* (283-300). Berlin: Springer.

Funke, J. & Reuschenbach, B. (2011). Einsatz technischer Mittel in der psychologischen Diagnostik. In L. Hornke, M. Amelang & M. Kersting (Hrsg.), *Leistungs-, Intelligenz- und Verhaltensdiagnostik* (=Enzyklopädie der Psychologie, Themenbereich B, Serie II, Band 3, 595-667). Göttingen: Hogrefe.

Frensch, P. A. & Sternberg, R. J. (1989) Expertise and intelligent thinking: When is it worse to know better? In R. J. Sternberg (Hrsg.), *Advances in the psychology of human intelligence* (157-188). Lawrence Erlbaum, Hillsdale, NJ.

Frey, A. (2006). Strukturierung und Methoden zur Erfassung von Kompetenz. *Bildung und Erziehung*, 59, 2, 125-166.

Fristoe, N. M. ; Salthouse, T. A. & Woodward, J. L. (1997). Examination of age-related deficits on the Wisconsin Card Sorting Test. *Neuropsychology*, 11, 428-436.

Gagné, R. M. (1965). *The conditions of learning and theory of instruction* (1st ed.). New York, NY: Holt, Rinehart & Winston.

- Gallagher, A. M., DeLisi, R., Holst, P.C., McGillicuddy-DeLisi, A. V., Morely, M. & Cahalan, C. (2000). Gender differences in advanced mathematical problem solving. *Journal of experimental Child Psychology*, 75, 165-190.
- Gallagher, A. (1997). Problem-based learning: Where did it come from, what does it do, and where is it going? *Journal for the Education of the Gifted*, 20, 4, 332-362.
- Gallagher, A., Stepien, W. J., & Rosenthal, H. (1992). The effects of problem-based learning on problem solving. *Gifted Child Quarterly*, 36, 4, 195-200.
- Gallagher, A., & Stepien, W. J. (1996). Content acquisition in problem-based learning: Depth versus breadth. *Journal for the Education of the Gifted*, 19, 3, 257-275;
- Miller, G. A., Galanter, E. & Pribram, K. H.(1960). *Plans and the structure of behavior*. New York: Henry Holt and Co.
- Gardner, H. (1991). *The unschooled mind: How children think and how schools should teach*. New York: Basic Book
- Gardner, H. (1999). *The disciplined mind: What all students should understand*. New York: Simon & Schuster.
- Genter, D. (1983). Structure-mapping. A theoretical framework for analogy. *Cognitive Science*, 7, 155-170.
- Geißler, K. A. & Orthey, M. F. (2002). Kompetenz: Ein Begriff für das verwertbare Ungefähre. In E. Nuissl, Ch. Schiersmann & H. Siebert (Hrsg.). *Literatur- und Forschungsreport Weiterbildung* (69-79). DIE, Bielefeld: Bertelsmann.
- Gigerenzer, G. (1988). Woher kommen Theorien über kognitive Prozesse? *Psychologische Rundschau*, 39, 91-100.
- Gilhooly, K. J. (1982). *Thinking: Directed, undirected, and creative*. London & New York: Academic Press.

Glug, I. (2009). Entwicklung und Validierung eines Multiple-Choice-Tests zur Erfassung prozessbezogener naturwissenschaftlicher Grundbildung. Dissertation. Zugriff am 19.01.2014 unter http://eldiss.uni-kiel.de/macau/receive/dissertation_diss_00003649

Gnahn, D. (2010). *Kompetenzen - Erwerb, Erfassung, Instrumente*. Bertelsmann: Bielefeld.

Gonon, Ph. (2002). Und am Schluss auch noch Latein? 30 Jahre Schlüsselqualifikationen. *Panorama Berufsberatung – Berufsbildung – Arbeitsmarkt*, 16, 6, 6-10.

Goldstone, R. L. & Son, J. Y. (2005). Similarity. In K. J. Holyoak & R. G. Morrison (Hrsg.), *The Cambridge Handbook of Thinking and Reasoning* (13-36), Cambridge: University Press.

Grabner, R. H., Stern, E. & Neubauer, A. C. (2003). When intelligence loses its impact: Neural efficiency during reasoning in a familiar area. *International Journal of Psychophysiology*, 49, 89-98.

Greeno, J. G. & Simon, H. A. (1988). Problem solving and reasoning. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey & R. D. Luce (Hrsg.), *Steven's Handbook of Experimental Psychology* (589-671). New York: Wiley & Sons.

Greeno, J. G., Collins, A., & Resnick, L. B. (1996). Cognition and learning. In D. C. Berliner & R. C. Calfee (Hrsg.), *Handbook of Educational Psychology* (15–46). New York: Macmillan.

Gruber, H. (1994). *Expertise. Modelle und empirische Untersuchungen*. Opladen: Westdeutscher Verlag.

Gruber, H. & Stamouli, E. (2009). Intelligenz und Vorwissen. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (27-47). Berlin, Heidelberg: Springer.

Gustaffson, J.-E. & Undheim, J. O. (1996). Individual Differences in cognitive function. In D. C. Berliner & R. C. Calfee (Hrsg.), *Handbook of Educational Psychology* (186-242). New York: Simon & Schuster Macmillan.

Grotzer, T. & Perkins, D. N. (2000). Teaching Intelligence: A Performance Conception. In R. J. Sternberg (Hrsg.), *Handbook of Intelligence* (492-518). Cambridge University Press.

- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E. & Tatham, R. L. (2006). *Multivariate Data Analysis*. Upper Saddle River, N.J.: Prentice-Hall.
- Hagmayer, Y. & Waldmann, M. R. (2006). Kausales Denken. In J. Funke (Hrsg.), *Enzyklopädie der Psychologie* (Denken und Problemlösen, Band II, 87-166). Göttingen: Hogrefe.
- Haladyna, T. M. (1994). *Developing and validating multiple-choice test items*. Hillsdale: Lawrence Erlbaum.
- Haladyna, T. M., Downing, M. & Rodriguez, M. C. (2002). A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. *Applied Measurement in Education*, 15, 3, 309-334.
- Haladyna, T. M. & Rodriguez, M. C. (2013). *Developing and Validating Testitems*. New York: Routledge.
- Hambrick, D. Z. & Engle, R. W. (2002). Effects of Domain Knowledge, Working Memory Capacity, and Age on Cognitive Performance: An Investigation of the Knowledge-Is-Power-Hypothese. *Cognitive Psychology*, 44, 339-387.
- Harp, F. & Mayer, R. E. (1997). The Role of Interest in Learning from Scientific Text and Illustrations: On the Distinction between Emotional Interest and Cognitive Interest. *Journal of Educational Psychology*, 89, 1, 92-102.
- Harp, F. & Mayer, R. E. (1998). How Seductive Details do their damage: A theory of Cognitive Interest in Science Learning. *Journal of Educational Psychology*, 90, 3, 414-434.
- Harskamp, E., Ding, N. & Suhre, C. (2008). Group composition and its effect on female and male problem-solving in science education. *Educational Research*, 50, 4, 307-318.
- Hartig, J. & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (127-143). Heidelberg: Springer.
- Hartig, J. & Jude, N. (2007). Empirische Erfassung von Kompetenzen. In J. Hartig & E. Klieme (Hrsg.), *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik. Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung* (17-36). Bonn, Berlin.

Hartig, J. & Klieme, E. (2007). *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik. Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung*. Bonn, Berlin.

Hawkins, J., Peas, R. D., Glick & Scribner, J. (1984). Merds that laugh don't like mushroom. Evidence for Deductive reasoning in Preschooler. *Developmental Psychology*, 20, 4, 584-594.

Healy, L., Pozzi, S. & Hoyles, C. (1995). Making Sense of Groups, Computers and Mathematics. *Cognition and Instruction*, 13, 4, 505-523.

Heigl, N. & Thomas, J. (2013). The effectiveness of epistemic beliefs and the moderating role of self-efficacy in the solving of cross-curricular problems. *Psychology Learning & Teaching*, 12(2), 126-135.

Heigl, N. & Thomas, J. (2012). Subjektive Überzeugungen als Komponenten fächerübergreifender Problemlösekompetenz - Ansatzpunkte für eine adaptive Lehre. In Mair M., Brezowar, G., Olsowski, G. & Zumbach, J. (Hrsg.). *Problem-Based Learning im Dialog. Anwendungsbeispiele und Forschungsergebnisse aus dem deutschsprachigen Raum*. Universitätsverlag facultas, Wien, S. 327- 339.

Hesse, F. W., Spies, K. & Lür, G. (1983). Einfluss motivationaler Faktoren auf das Problemlöseverhalten im Umgang mit komplexen Problemen. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 30, 3, 400-424.

Hmelo, C. E., Gotterer, G. & Bransford, J. D. (1997). A theory-driven approach to assessing the cognitive effects of problem-based learning. *Instructional Science*, 25, 387-408.

Hmelo-Silver, C. E. (2004). Problem-Based Learning: What and How Do Students Learn? *Educational Psychology Review*, 16, 3, 235-266.

Hochschulrektorenkonferenz, Kultusministerkonferenz, Bundesministerium für Bildung und Forschung (2005). Qualifikationsrahmen für Deutsche Hochschulabschlüsse. Zugriff am 12.01.2014 unter <http://www.hrk.de/de/download/dateien/QRfinal2005.pdf>

Hofer, B. (2001). Personal Epistemology Research: Implications for Learning and Teaching. *Journal of Educational Psychology Review*, 13, 4, 353- 383.

- Hofer, B. K. (2002). Personal Epistemology as an Psychological and Educational Construct: An Introduction. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal Epistemology: The Psychology of Beliefs about knowledge and knowing* (3-14). New Jersey: Erlbaum.
- Hofer, B. (2004). Introduction: Paradigmatic Approaches to Personal Epistemology. *Educational Psychologist*, 39, 1, 1-3.
- Hofer B. (2008). Personal Epistemology and Culture. In M. Khine (Hrsg.), *Knowing, Knowledge, and Beliefs: Epistemological Studies across Diverse Cultures* (3-22), Springer.
- Holyoak, K. J. & Morrison R. G. (2005). *The Cambridge Handbook of Thinking and Reasoning*. Cambridge: University Press.
- Holyoak, M. (2005). Analogy. In K. J. Holyoak, R. G. Morrison (Hrsg.), *The Cambridge Handbook of Thinking and Reasoning* (117-142), Cambridge: University Press.
- Holyoak, M. & Koh, K (1987). Surface and structural similarity in analogical transfer. *Memory and Cognition*, 15, 332-340.
- Hood, A. & Deopere, D. L. (2002). The Relationship of Cognitive Development to Age, When Education and Intelligence are controlled for. *Journal of Adult Development*, 9, 3, 229-234.
- Hu, Li-tze & Bentler P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives. *Structural Equation Modeling*, 6, 1-55.
- Hussy, W. (1984). Zum Begriff der Problemschwierigkeit beim komplexen Problemlösen. In R. Ferstl, H. Gräser & R. Scheller (Hrsg.). *Trierer Psychologische Berichte*, Band 11, Heft 4.
- Hunt, E. (1994). Problem Solving. In R. J. Sternberg (Hrsg.), *Thinking and Problem Solving* (215-234). London, Kalifornien: Academic Press.
- Huttenlocher, J. (1968). Constructing spatial images: A strategy in reasoning. *Psychological Review*, 75, 550-560.
- Isen, A. M. (2001). An Influence of Positive Affect on Decision Making in Complex Situations: Theoretical Issues with Practical Implications. *Journal of Consumer Psychology*, 11, 2, 75-85.

- James, W. (1890). *The principles of psychology*. Vol I. New York: Henry Holt and Co.
- Jaspers, K. (1923). *Die Idee der Universität*. Berlin, Heidelberg, New York: Springer.
- Jelsma, O., Merrienboer, J. J. G. Van & Bijlstra J. P. (1990). The ADAPT design model: towards instructional control of transfer. *Instructional Science*, 19, 89-120.
- Jerusalem, M. & Satow, L. (1999). Schulbezogene Selbstwirksamkeitserwartung. In R. Schwarzer & M. Jerusalem (Hrsg.), *Skalen zur Erfassung von Lehrer- und Schülermerkmalen Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen*. Zugriff am 01.07.2013 unter http://userpage.fu-berlin.de/~health/self/skalendoku_selbstwirksame_schulen.pdf
- Johnson-Laird, P. N. & Wason, P.C. (1970). A Theoretical Analysis of Insight Into a Reasoning Task. *Cognitive Psychology*, 1, 134-148.
- Johnson-Laird, P. N. & Byrne, R. J. M. (1989). Only reasoning. *Journal of Memory and Language*, 28, 3, 313-330.
- Johnson-Laird, P. (1995). Mental Models, Deductive Reasoning and the Brain. In M. Gazzaniga (Hrsg.), *The Cognitive Neurosciences*. Cambridge, Bradford Book: MIT Press.
- Jonassen, D. H. (1997). Instructional Design Models for Well-structured and Ill-structured Problem-Solving Learning Outcome Educational Technology. *Research and Development*, 45, 1, 65-95.
- Jonassen, D. H. (2000). Toward a Design Theory of Problem Solving. Educational Technology. *Research and Development*, 48, 4, 63-85.
- Jones, H. & Davidson, R. A. (2007). Measuring the Problem-Solving Abilities of Accounting and Other Business Students: A Comparison and Evaluation of Three Method Accounting Education. *International Journal*, 16, 1, 65-79.
- Kaartinen- Koutaniemi, M. & Lindblom-Ylänne, S. (2008). Personal epistemology of psychology, theology and pharmacy students: a comparative study. *Studies in Higher Education*, 33, 2, 179-191.

Kahnemann, D., Slovic, P. & Tversky, A. (1982). *Judgement under uncertainty. Heuristics and biases*. Cambridge: Cambridge Press.

Kendler, H. H. & Kendler, T. (1962). Vertical and horizontal processes in problem solving. *Psychological Review*, 69, 1, 1-16.

Kennedy, C. A. (2005). *The BEAR Assessment System: A Brief Summary for the Classroom Context* (BEAR Technical Report Series, No. 2005-03-01). Berkeley Evaluation & Assessment Research (BEAR) Center: University of California, Berkeley.

Kennedy, C. A. & Wilson, M. (2007). *Using Progress Variables to Interpret Student Achievement and Progress* (BEAR Technical Report No. 2006-12-01). Berkeley Evaluation & Assessment Research (BEAR) Center: University of California, Berkeley.

Kemp, I. J. & Seagraves, L. (1995). Transferable skills – can higher education deliver? *Studies in Higher Education*, 20, 3, 315-328.

Kintsch, W. (1998). *Comprehension. A paradigm for cognition*. Cambridge, UK: Cambridge University Press.

Kirsh, D. (2009). Problem Solving and Situated Cognition. In P. Robbins & M. Aydede (Hrsg.), *Situated Cognition* (264-306). New York: Cambridge Press.

Knauff, M. (2006). Deduktion und logisches Denken. In J. Funke (Hrsg.), *Enzyklopädie der Psychologie* (Themenbereich C, Bd. 8, Denken und Problemlösen, 167-264). Göttingen: Hogrefe.

Knight, P. T. & Yorke, M. (2003). Employability and Good Learning in Higher Education. *Teaching in Higher Education*, 8, 1, 3-16.

Knoblich, G. (2002). Problemlösen und logisches Schließen. In J. Müsseler & W. Prinz (Hrsg.), *Allgemeine Psychologie* (644-701). Heidelberg: Spektrum.

Klauer, K. J. (2001). *Handbuch kognitives Training*. Göttingen, Bern Toronto, Seattle: Hogrefe.

Klep, J. & Haperen T. v. (2004). Competence: conceptual exploration. In Consortium of Institutions for Development in Education in Europe, Netherlands Institute for Curriculum

Develoment (Hrsg.), *The integrated person. How curriculum development relates to nee perspective* (25-34). Enschede.

Klemp, G.O., Jr. & McClelland, D.C. (1986). What characterizes intelligent functioning among senior managers? In R. J. Sternberg & R. K. Wagner (Hrsg.), *Practical intelligence: Nature and origins of competence in the everyday world* (31-50). Cambridge: Cambridge University Press.

Klieme, E. (2004). Was sind Kompetenzen und wie lassen sie sich messen? *Pädagogik*, 6, 10-13.

Klieme, E. & Hartig, J. (2007a). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In M. Prenzel, I. Gogolin & H.-H. Krüger (Hrsg.), *Kompetenzdiagnostik* (Zeitschrift für Erziehungswissenschaft, Sonderheft 8, 11-29). Wiesbaden: V.

Klieme, E., Maag-Merki, K. & Hartig, J. (2007b). Kompetenzbegriff und Bedeutung von Kompetenzen im Bildungswesen. In J. Hartig & E. Klieme (Hrsg.), *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik. Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung* (5-15). Bonn, Berlin.

Klieme, E., Stanat, P. & Artelt, C. (2001b). Fächerübergreifende Kompetenzen: Konzepte und Indikatoren. In F. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (203-218). Weinheim, Basel: Beltz.

Klieme, E., Neubrand, M. & Lüdtke, O. (2001). Mathematische Grundbildung: Testkonzeption und Ergebnisse. In Deutsches PISA- Konsortium (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (139–190). Opladen: Leske & Budrich.

Klieme, E., Funke, J., Leutner, D., Reimann, P. & Wirth, J. (2001d). Problemlösen als fächerübergreifende Kompetenz. Konzeption und erste Resultate aus einer Schulleistungsstudie. *Zeitschrift für Pädagogik*, 47, 2, 179-200.

Klieme, E., Leutner, D. & Wirth, J. (2005). *Problemlösekompetenz von Schülerinnen und Schülern. Diagnostische Ansätze, theoretische Grundlagen und empirische Befunde der deutschen PISA-2000-Studie*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Klix, F. (1971). *Information und Verhalten. Kybernetische Aspekte der organismischen Informationsverarbeitung*. Bern: Huber.

- Koepfen, K., Hartig, J., Klieme, E. & Leutner, D. (2008). Current Issues in Competence Modeling and Assessment, *Zeitschrift für Psychologie*, 216, 2, 61–73.
- Krampen, G. (1993). Diagnostik der Kreativität. In G. Trost, K. Ingenkamp & R. S. Jäger (Hrsg.), *Tests und Trend* (11–39). Beltz, Weinheim.
- Krapp, A. & Ryan, R. M. (2002). Selbstwirksamkeit und Lernmotivation. In M. Jerusalem & D. Hopf (Hrsg.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen* (Zeitschrift für Pädagogik 44. Beiheft, 54-82).
- Krech, D., Crutchfield, R. & Livson, N. (1974). *Elements of psychology*. Oxford, England: Alfred A. Knopf.
- Krosnick, J. A. (1999). Survey Research. *Annual Review of Psychology*, 50, 537-567.
- Kutschera, F. v., & Bretkopf, A. (1985). *Einführung in die moderne Logik*. Freiburg: Karl Alber.
- Koschmann, T., Kelson, A.C., Feltovich, P. J. & Barrows, H. (1996). Computer-supported problem-based learning: A principled approach to the use of computers in collaborative learning. In T. Koschmann (Hrsg.), *CSCL: Theory and practice of an emerging paradigm*. Mahwah, NJ: Lawrence Erlbaum.
- Kotovsky, K. & Simon, H. A. (1990). What makes some problems really hard: Explorations in the problem space of difficulty. *Cognitive Psychology*, 22, 143-183.
- Kotovsky, K. (2003). Problem Solving – Large/Small, Hard/Easy, Conscious/Nonconscious, Problem-Space/Problem-Solver. The issue of Dichotomization. In J. E. Davidson & R. J. Sternberg (Hrsg.), *The Psychology of Problem Solving* (373-383). Cambridge: University Press.
- Larkin, J. H., McDermott, J., Simon, D. P. & Simon, H. A. (1980a). Expert and Novice Performance in Solving Physics Problems. *Science*, 208, 1335-1342.
- Larkin, J. H., McDermott, J., Simon, D. P. & Simon, H. A. (1980b). Models of competence in solving physics problem. *Cognitive Science*, 4, 317–345.
- Lefford, A. (1946). The influence of emotional subject matter on logical reasoning. *Journal of General Psychology*, 34, 127-151.

- Lemaire, P. & Siegler, R. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General*, 124, 83-97.
- Leuders, T. & Sodian, B. (2013). Inwiefern sind Kompetenzmodelle dazu geeignet kognitive Prozesse von Lernenden zu beschreiben? *Zeitschrift für Erziehungswissenschaft*, 16, 27–33.
- Leutner, D. (2002). The fuzzy relationship of intelligence and problem solving in computer simulations. *Computers in Human Behavior*, 18, 6, 685–697.
- Leutner, D., Klieme, E., Meyer, K. & Wirth, J. (2004). Problemlösen. In PISA-Konsortium Deutschland (Hrsg.), *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs* (147-175). Münster: Waxmann.
- Leutner, D., Wirth, J., Klieme, E. & Funke, J.(2005). Ansätze der Operationalisierung und deren Erprobung im Feldtest zu PISA 2000. In E. Klieme & D. Leutner & J. Wirth (Hrsg.), *Problemlösekompetenz von Schülerinnen und Schülern. Diagnostische Ansätze, theoretische Grundlagen und empirische Befunde der deutschen PISA-2000-Studie* (21-36). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Lienert, G. A. & Ratz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Psychologie Verlags Union.
- Lim, J., Reiser, R. A. & Olina, Z. (2009). The effects of part-task and whole-task instructional approaches on acquisition and transfer of complex cognitive skill. *Educational Technology Research and Development*, 57, 61-77.
- Lipshitz, R., Levy, D. L. & Orchen, K. (2006). Is this problem likely to be solved? A cognitive schema of effective problem solving. *Thinking & Reasoning*, 12, 4, 413-430.
- Lizzio, A., Wilson, K. & Simons, R. (2002). University Students Perceptions of the Learning Environment and Academic Outcomes: implications for theory and practice. *Studies in Higher Education*, 27, 1, 27-52.
- Lompscher, J. & Mandl, H. (1996). *Lehr- und Lernprobleme im Studium*. Bern: Huber.

- Lüer, G. & Spada, H. (1990). Denken und Problemlösen. In H. Spada (Hrsg.), *Lehrbuch Allgemeine Psychologie* (189-280). Bern: Huber.
- Luchins, A. (1942). Mechanization in problem solving: The effect of *Einstellung*. *Psychological Monographs*, 54, 6, i-95.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. *Psychologische Rundschau*, 58, 2, 103-117.
- Maag-Merki, K., Grob, U. (2001). *Überfachliche Kompetenzen. Theoretische Grundlegung und empirische Erprobung eines Indikatorensystems*. Bern: Lang.
- Mähler, C. & Stern, E. (2006). Transfer. In D. H. Rost (Hrsg.), *Handwörterbuch: Pädagogische Psychologie* (3. überarbeitete und erweiterte Aufl., 782-793). Weinheim: Beltz.
- Manninen, J. & Hobrough, J. (2000). Skills gaps and overflows? *Industry and Higher Education*, 14, 1, 51-57.
- Markovitz, H. & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory & Cognition*, 17, 11-17.
- Marsh, H. W. & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, 81, 59-77.
- Mayer, R. E. (1979). *Denken und Problemlösen. Eine Einführung in menschliches Denken und Lernen*. Berlin: Springer.
- Mayer, R. E. (1992). *Thinking, problem solving, cognition*. New York: W H Freeman/Times Books/ Henry Holt & Co.
- Mayer, R. E. & Wittrock, A. C. (1996): Problem Solving and Transfer. In D. C. Berliner, R. C. Calfee (Hrsg.), *Handbook of educational Psychology* (47-62). New York: Simon & Schuster McMillan.
- Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem Solving. *Instructional Science*, 26, 49-63.

- Mayer, R. E. (2002). A taxonomy for computer-based assessment of problem solving. *Computers in Human Behavior*, 18, 6, 623–632.
- Mayer, R. E. (2003). What causes individual differences in cognitive performance? In R. J. Sternberg & E. L. Grigorenko (Hrsg.), *The psychology of abilities, competencies, and expertise* (263-274). Cambridge University Press.
- Mayer, R. E. (2010). Problem solving and Reasoning. *Learning and Cognition*, 273-278.
- McPherson, K. (2002). Problem-solving Ability and Cognitive Maturity in Undergraduate Students. *Assessment & Evaluation in Higher Education*, 27, 1, 5-22.
- Mertens, K. (1974). Schlüsselqualifikationen. *MittAB*, 1, 36-43.
- Mertens, D. (1988): Das Konzept der Schlüsselqualifikationen als Flexibilisierungsinstrument. *Report*, 22, 33-46.
- Molden, D.C. & Higgins, E. T. (2005). Motivated Thinking. In K. J. Holyoak & R. G. Morrison (Hrsg.), *The Cambridge Handbook of Thinking and Reasoning* (295-320). Cambridge: University Press.
- Moosbrugger, H. & Kelava, A. (2008). *Testtheorie und Testkonstruktion*. Heidelberg: Springer.
- Moschner, B & Gruber, H. & die Studienstiftungsarbeitsgruppe EPI (2005). *Epistemologische Überzeugungen* (Forschungsbericht 18). Regensburg: Universität Regensburg, Lehrstuhl für Lehr-Lern-Forschung.
- 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, 3-54.
- Multrus, F., Ramm, M. & Bargel, T. (2011). *Studiensituation und studentische Orientierungen. 11. Studierendensurvey an Universitäten und Fachhochschulen*. Bundesministerium für Bildung und Forschung (BMBF). Bonn, Berlin.
- Mumford, M. D., Reiter-Palmon, R. & Redmond, M. R. (1994). Problem construction and cognition: Applying problem representations in ill-defined domain. In M. A. Runco (Hrsg.),

- Problem finding, problem solving, and creativity* (3-39). Norwood, New Jersey: Ablex Publishing Corporation.
- Murray, K. & Macdonald, R. (1997). The disjunction between lecturers' conceptions of teaching and their claimed educational practice. *Higher Education*, 33, 3, 331-349.
- Musch, J. & Lieberei, W. (1997). *Eine Auswertungsobjektive Postkorbübung für Assessmentcenter*. Berichte aus dem Psychologischen Institut der Universität Bonn, Band 23, Heft 1.
- Myers, C. & Conner, M. (1992). Age Differences in Skill Acquisition and Transfer in an implicit Learning Paradigm. *Applied Cognitive Psychology*, 6, 429-442.
- Nabi, G. R. & Bagley, D. (1999). Graduates perceptions of transferable personal skills and future career preparation in the UK. *Education and Training*, 41, 4, 184-193.
- Nährer, W. (1988). Problemkonstruieren statt Problemlösen. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 35, 441-450.
- Neisser, U. (1996). Intelligence: Knowns and Unknowns. *American Psychologist*, 51, 2, 77-101.
- Neumann, R., Parry, S. & Becher, T. (2002). Teaching and learning in their disciplinary context. A conceptual analysis. *Studies in Higher Education*, 27, 405-417.
- Newell, A., Shaw, J. C. & Simon, H. A. (1958a). Chess-Playing Programs and the Problem of Complexity, *IBM Journal of Research and Development*, 2, 4, 320 – 335.
- Newell, A., Shaw, J. C. & Simon, H. A. (1958b). Elements of a Theory of human Problem Solving. *Psychological Review*, 65, 3, 151-166.
- Newell, A. & Simon, H. A. (1961a). Computer simulation of thinking. *Science*, 134, 2011-2017.
- Newell, A. & Simon, H. A. (1961b). GPS, A Program that Simulates Human Thought. In H. Billing (Hrsg.), *Lernende Automaten*. München: Oldenbourg, 109-124.
- Nickerson, R. (1994). The Teaching of Thinking and Problem Solving. In R. J. Sternberg (Hrsg.), *Thinking and Problem Solving* (409-449). London, Kalifornien: Academic Press.

- Nickerson, R. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, 2, 2, 175-220.
- Nickerson, R., Perkins, D. N. & Smith, E. E. (1985). *The Teaching of Thinking*. Hillsdale, N.J.: L. Erlbaum Associates.
- Niedelman, M. (1991). Problem Solving and Transfer. *Journal of Learning Disabilities*, 24, 6, 322-329.
- Nokey, T. (2009). Mechanisms of knowledge transfer. *Thinking and Reasoning*, 15, 1, 1-36.
- Novick, L. R. & Bassok, M. (2005). Problem Solving. In K. J. Holyoak & R. G. Morrison (Hrsg.), *The Cambridge Handbook of Thinking and Reasoning* (321-350). Cambridge: University Press.
- Novick, L. R. (1995). Some determinants of successful analogical transfer in the solution of algebra word problems. *Thinking and Reasoning*, 1, 5-30.
- Norton, L., Richardson, J. T. E., Hartley, J., Newstead, S. & Mayes, J. (2005). Teachers' beliefs and intentions concerning teaching in higher education. *Higher Education*, 50, 537-571.
- Norman, G.R., & Schmidt, H.G. (1992). The psychological basis of problem-based learning: A review of the evidence. *Academic Medicine*, 67, 557-565.
- Nowakowski, A., Vervecken, D., Braun, E. & Hannover, B. (2012). Was Hochschuldozierende aus Lehrevaluationen lernen können. *Zeitschrift für Erziehungswissenschaften*, 15, 253-271.
- OECD Testaufgaben PISA. Zugriff am 12.01.2014 unter <http://www.oecd.org/berlin/themen/pisa-internationaleschulleistungstudiederocd.htm>.
- OECD (2004). *Problemsolving for tomorrow's world. First measurements of cross-curricular competencies from PISA 2003*. Paris: OECD.
- Ohlsson, St. (2012). The Problems with Problem Solving: Reflections on the Rise, Current Status, and Possible Future of a Cognitive Research Paradigm. *The Journal of Problem Solving*, 5, 1, 101-128.
- Ortega y Gasset, J. (1944). *Mission of the university*. Princeton, NJ: Princeton University Press.

Orthey, F. M. (2002). Der Trend zur Kompetenz. *Supervision*, 1, 7-14.

Pajares, M. F. (1992). Teachers beliefs and educational research: Cleaning up a messy construct. *Review of educational Research*, 62, 307-332.

Pajares, F. & Miller, M. D. (1994). Role of Self-Efficacy and Self-Concept Beliefs in Mathematical Problem Solving: A Path Analysis. *Journal of Educational Psychology*, 86, 2, 193-203.

Patel, V. L., & Groen, G.J. (1991). The general and specific nature of medical expertise: A critical look. In K.A. Ericsson & J. Smith (Hrsg.), *Toward a General Theory of Expertise: Prospects and Limits* (93-125). New York: Cambridge University Press.

Patel, V. K., Groen, G. J. & Norman, G.R. (1993). Effects of conventional and problem-based medical curricula on problem solving. *Academic Medicine*, 66, 380-389.

Pearl, J. (2000). *Causation: Models, reasoning and inference*. Cambridge, MA: Cambridge University Press.

Pellegrino, J. W., Chudowsky, N. & Glaser, R. (2001). *Knowing what students know. The science and design of Educational Assessment*. Washington: National Academy Press.

Perkins, D. N. & Simmons, R. (1988). Patterns of Misunderstanding: An Integrative Model for Science, Math, and Programming. *Review of Educational Research*, 58, 3, 303-326.

Perkins, D. N. & Salomon, G. (1987). Transfer and Teaching of Thinking. In D. N. Perkins, J. Lochhead & J. Bishop (Hrsg.), *Thinking: The second international conference* (285-303). Erlbaum, Hillsdale, New Jersey.

Perkins, D. N. & Salomon, G. (1989). Are cognitive skills context-bound? *Educational Researcher*, 18, 1, 16-25.

Perkins, D. & Salomon, G. (1992). *Transfer of Learning*. Oxford, England: Pergamon Press.

Perkins, D. (1993). Teaching for Understanding. *American Educator*, 17, 3, 28-35.

Perkins, D. & Salomon, G. (o. J.). The Science and Art of Transfer. Zugriff am 12.01.2014 unter <http://learnweb.harvard.edu/alps/thinking/docs/trancost.pdf>

- Peschar, J. L. (2004). Cross-curricular competencies. Developments in a new area of education outcome indicators. In E. Owen (Hrsg.), *International Education Outcome Indicators* (79-109). Paris: OECD.
- Pintrich, P.R. & De Groot, E.V. (1990). Motivation and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82, 33-40.
- Polya, G. (1985, 2. Aufl.). *How to solve it*. Princeton: University Press.
- Poropat, A. E. (2011). The Eysenckian personality factors and their correlations with academic performance. *British Journal of Educational Psychology*, 81, 41-58.
- Pozzi, St., Healy, L. & Hoyle, C. (1993). Learning and Interaction in Groups with Computers: When do ability and gender matter? *Social Development*, 2, 3, 222-242.
- Pretz, J. E., Naples, A.J. & Sternberg, R. J. (2003). Recognizing, Defining, and Representing Problems. In J. E. Davidson & R. J. Sternberg (Hrsg.), *The Psychology of Problem Solving* (3-30). Cambridge: University Press.
- Prenzel, M., Rost, J., Senkbeil, M., Häußler, P. & Klopp, A. (2001). Naturwissenschaftliche Grundbildung: Testkonzeption und Ergebnisse. In Deutsches PISA-Konsortium (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (192-248). Opladen: Leske & Budrich.
- Priemer, B. (2006). Deutschsprachige Verfahren der Erfassung von epistemologischen Überzeugungen. *Zeitschrift für Didaktik der Naturwissenschaften*, 12, 159-175.
- Prosser, M., Trigwell, K. (1999). *Understanding learning and teaching: the experience in higher education*. Buckingham: The Society for Research into Higher Education & Open University Press.
- Ramm, M. (2001). *Berufswahl, Berufsperspektiven und Existenzgründungen. Berufliche Orientierungen von Studierenden*. Bonn, Berlin: Bundesministerium für Bildung und Forschung (BMBF).
- Raviv, A, Bar-Tal, D., Raviv, A., Biran, B. & SELA, Z. (2003). Teachers epistemic authority: Perceptions of students and teachers. *Social Psychology of Education*, 6, 17-42.

- Rasch, B., Frieze, M., Hofmann, W. J. & Naumann, E. (2010). *Quantitative Methoden*. (Band 1, 3. Auflage). Heidelberg: Springer.
- Reeff, J.-P., Zabal, A. & Blech, Ch. (2006). The Assessment of Problem-Solving Competencies. A draft version of a general framework. Deutsches Institut für Erwachsenenbildung. Zugriff am 12.01.2014 unter http://www.die-bonn.de/esprid/dokumente/doc-2006/reeff06_01.pdf
- Reetz, L. (2006) Schlüsselqualifikationen aus bildungstheoretischer Sicht. In R. Arnold & H. J. Müller (Hrsg.), *Kompetenzentwicklung durch Schlüsselqualifikations-Förderung* (39–54). Baltmannsweiler: Schneider.
- Reitman, W. R. (1965). *Cognition and Thought*. New York: John Wiley & Sons.
- Renninger, K.A., Hidi, S. & Krapp, A. (1992). *The Role of Interest in Learning and Development*. Hillsdale, NJ: Erlbaum.
- Renkl, A. (1996). Träges Wissen: Wenn Erlerntes nicht genutzt wird. *Psychologische Rundschau*, 47, 78-92.
- Renkl, A. & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist*, 38, 15-22.
- Reinfeldt, F. & Frings, C. (2003): Absolventenbefragungen im Kontext von Hochschulevaluation - Forschungsstand und Perspektiven. *Zeitschrift für Evaluation*, 2, 279-294.
- Resnick, L. B. (1987). Constructing knowledge in school. In L. Liben (Hrsg.), *Development and learning: Conflict or congruence?* (19-50). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Resnick, L. B. (1992). From protoquantities to operators: Building mathematical competence on a foundation of everyday knowledge. In G. Leinhardt, R. Putnam & R.A. Hatrup (Hrsg.), *Analysis of arithmetic for mathematics teaching* (373-429). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Reuschenbach, B. & Funke, J. (2000). Management-Fallstudien (MFA). In E. Fay (Hrsg.), *Tests unter der Lupe III* (83-112). Lengerich: Pabst Science Publisher.

- Ritchardt, R. & Perkins, D. N. (2005). Learning to Think: The Challenges of Teaching Thinking. In K. J. Holyoack, R. G. Morrisson (Hrsg.), *The Cambridge Handbook of Thinking and Reasoning* (775-802). Cambridge: University Press.
- Rips, L. J. (1994). Deduction and its cognitive basis. In R.S. Sternberg (Hrsg.), *Thinking and Problem Solving* (150-180). London, Kalifornien: Academic Press.
- Rips, L. J. (2008). Causal Thinking. In J. E. Adler & L. J. Rips (Hrsg.), *Reasoning: Studies of Human Inference and Its Foundation*. Cambridge, UK: Cambridge University Press.
- Robbins, P. & Aydede, M. (2009) *Situated Cognition*. New York: Cambridge Press.
- Rost, J. (2004). Psychometrische Modelle zur Überprüfung von Bildungsstandards anhand von Kompetenzmodellen. *Zeitschrift für Pädagogik*, 50, 5, 662-678.
- Royer, J. M., Cisero, Ch. A. & Carlo, M. (1993). Techniques and Procedures for Assessing Cognitive Skills. *Review of Educational Research*, 63, 2, 201-243.
- Rumelhart, D. E., Lindsay, P. H. & Norman, D. A. (1972). A process model for long-term memory. In E. Tulving & W. Donaldson (Hrsg.), *Organization of memory* (198-246). New York: Academic Press.
- Rychen, D. & Salganik, L. H. (2003). *Key Competencies for a Successful Life and a Well-Functioning Society*. Göttingen: Hogrefe, Huber.
- Sandoval, W. A. & Morrisson, K. (2003). High school students' ideas about theories and theory change after a biological inquiry unit. *Journal of Research in Science Teaching*, 40, 4, 369-392.
- Salomon G. & Perkins, D. N. (1989). Rocky Roads to Transfer: Rethinking Mechanism of a Neglected Phenomenon. *Educational Psychologist*, 24, 2, 113-142.
- Salthouse, T. A. (1989). Aging and skilled performance. In A. M. Colley & J. R. Beech (Hrsg.), *Acquisition and performance of Cognitive Skill* (248-264). Wiley.
- Salthouse, T. A. (1991). *Theoretical Perspectives on cognitive aging*. Hillsdale, NJ: Erlbaum.
- Salthouse, T. A. (1993). Influence of working memory on adult age differences in matrix reasoning. *British Journal of Psychology*, 84, 171-199.

Salthouse, T. A. (1996). The processing speed theory of adult age differences in cognition. *Psychological review*, 193, 403-428.

Salthouse, T. A. (2005). Effects of Aging on Reasoning. In K. J. Holyoak & R. Morrison (Hrsg.), *Cambridge handbook of thinking and reasoning* (589-606). New York: Cambridge University Press.

Saugstadt, P. & Raaheim, K. (1968). Problem-Solving, Past Experience and Availability of Function In P.C. Wason & P. N. Johnson-Laird (Hrsg.), *Thinking and Reasoning, Selected Readings*. Penguin Modern Psychology.

Schecker H. & Parchmann, I. (2006). Modellierung naturwissenschaftlicher Kompetenz. *Zeitschrift für Didaktik der Naturwissenschaften*, 12, 45-66.

Scheiter, K. (2003). Sequenz- und Sequenzierungseffekte bei der Bearbeitung voraussetzungsreicher Aufgaben. Dissertation. Zugriff am 12.01.2014 unter <http://nbn-resolving.de/urn:nbn:de:bsz:21-opus-10972>

Schiano, D. J., Cooper, L. A., Glaser, R. & Zang, H. C. (1989). Highs are to Lows as Experts are to Novices: Individual Differences in the Representation and Solution of Standardized Figural Analogies. *Human Performance*, 2, 4, 225-248.

Schiefele, U (1988). Der Einfluss von Interesse auf Umfang, Inhalt und Struktur studienbezogenen Wissens. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, XX, 4, 356-370.

Schiefele, U. (1992). Topic interest and level of text comprehension. In K. A. Renninger, S. Hidi & A. Krapp (Hrsg.), *The Role of Interest in Learning and Development* (151-182). Hillsdale, NJ: Erlbaum.

Schiefele, U., Krapp, A. & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi & A. Krapp (Hrsg.), *The Role of Interest in Learning and Development* (183–212). Hillsdale, NJ: Erlbaum.

Schiefele, U., Krapp, A. & Schreyer, I. (1993). Metaanalyse des Zusammenhangs von Interesse und Leistung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, XXV, 2, 120-148.

- Schiefele, U. Krapp, A. & Winterler, A. (1993). Der "Fragebogen zum Studieninteresse" (FSI). *Diagnostica*, 39, 4, 335-351.
- Schneider, W. & Bjorklund, D. F. (1992). Expertise, aptitude, and strategic remembering. *Child Development*, 63, 461-473.
- Schneider, W. (1985). Training high performance skills: Fallacies and guidelines. *Human Factors*, 27, 3, 285-300.
- Schnotz, W. & Bannert, M. (1999). Einflüsse der Visualisierungsform auf die Konstruktion mentaler Modelle beim Text- und Bildverstehen. *Zeitschrift für Experimentelle Psychologie*, 46, 3, 217-236.
- Schnotz, W., Vosniadou, S. & Carretero, M. (1999). *New Perspectives in Conceptual Change Research*, Elsevier.
- Schoenfeld, A. (1983). Beyond the purely cognitive: Belief systems, social cognitions, and metacognitions as driving forces in intellectual performance. *Cognitive Science*, 7, 329-363.
- Schommer-Aikins, M., Duell, O. K. & Hutter, R. (2005). Epistemological Beliefs, Mathematical Problem-Solving Beliefs, and Academic Performance of Middle School Students. *The Elementary School Journal*, 105, 3, 289-304.
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82, 498-504.
- Schraw, G., Dunkle, M. E. & Bedixen L. D. (1995). Cognitive Processes in Well-defined and Ill-defined Problem Solving. *Applied Cognitive Psychology*, 9, 523-538.
- Schwarz, N. & Skurnik, I. (2003). Feeling and Thinking: Implications for Problem Solving. In J. E. Davidson & R. J. Sternberg (Hrsg.), *The Psychology of Problem Solving* (263-290). Cambridge: University Press.
- Schwarzer, R. (1993). *Streß, Angst und Handlungsregulation*. Stuttgart: Kohlhammer.
- Schwarzer, R. (1995). *Entwicklungskrisen durch Selbstregulation meistern*. Zugriff am 10.05.2013 unter <http://paedpsych.jk.uni-linz.ac.at:4711/LEHRTEXTE/Schwarzer.html>

- Schwarzer, R. & Jerusalem, M. (1999). *Skalen zur Erfassung von Lehrer- und Schülermerkmalen Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen*. Zugriff am 10.5.2013 unter http://userpage.fu-berlin.de/~health/self/skalendoku_selbstwirksame_schulen.pdf
- Schwarzer, R. & Jerusalem, M. (2002). Das Konzept der Selbstwirksamkeit. In M. Jerusalem & D. Hopf (Hrsg.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen* (28-53). Weinheim: Beltz.
- Schunk, D. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, 26, 207- 231.
- Schunk, D. H. & Hanson, A. R. (1985). Peer models: Influences on children's self-efficacy and achievement. *Journal of Educational Psychology*, 77, 313–322.
- Segers, M., Dochy, F. & De Corte, E. (1999). Assessment practices and students' knowledge profiles in a problem-based curriculum. *Learning Environments Research*, 2, 191-213.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57, 1-22.
- Shavelson, R. J. & Stern, P. (1981). Research on Teachers' pedagogical thoughts, judgements, decisions and behavior. *Review of Educational research*, 51, 455-498.
- Schunn, C. D. & Anderson, J. R. (1999). The generality / specificity of expertise in scientific reasoning. *Cognitive Science*, 23, 337-370.
- Simeaner, H., Ramm, M., Kolbert-Ramm, C. (2010). *Datenalmanach Studierendensurvey 1993 - 2010. Studiensituation und Studierende an Universitäten und Fachhochschulen* (Heft 59). Konstanz, Arbeitsgruppe Hochschulforschung, Universität Konstanz.
- Simon, H. A. (1975). The Functional Equivalence of Problem Solving Skills. *Cognitive Psychology*, 7, 268-288.
- Simon, H. A. & Barenfeld, M. (1969). Information-Processing Analysis Of Perceptual Processes In Problem Solving. *Psychological review*, 67, 5, 473-483.

Simon, D. P. & Simon, H. A. (1978). Individual differences in solving physics problems. In R. Siegler (Hrsg.), *Children's thinking: What develops?* (325-348). Hillsdale, NJ: Erlbaum.

Simon, H. A. & Newell, A. (1970). *Human Problem Solving*. Englewood Cliffs, N. J.: Prentice-Hall.

Simon, H. A. (1990). Invariants of Human Behavior. *Annual Review of Psychology*, 41, 1-19.

Simonton, D. K. (2003). Expertise, Competence, and Creative Ability: The Perplexing Complexity. In R. J. Sternberg & E. L. Grigorenko (Hrsg.), *The psychology of abilities, competencies, and expertise* (213-239). Cambridge University Press.

Singely, K. & Anderson, J. R. (1989). *The Transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Slooman, A. & Lagnado, D. A. (2005). The Problem of Induction. In K. J. Holyoak & R. G. Morrison (Hrsg.), *The Cambridge Handbook of Thinking and Reasoning* (95-116), Cambridge: University Press.

Snow, R. E., Corno, L. & Jackson III, D. (1996). Individual Differences in affective and conative functions. In D. C. Berliner & R. C. Calfee (1996). *Handbook of educational Psychology* (243-310), New York: Simon & Schuster McMillan.

Star, C. & Hammer (2008). Teaching generic skills: eroding the higher purpose of universities, or an opportunity for renewal? *Oxford Review of Education*, 34, 2, 237-251.

Stanovich, K. E. (2003). The Fundamental Computational Biases of Human Cognition. In J. E. Davidson & R. J. Sternberg (Hrsg.), *The Psychology of Problem Solving* (291-342). Cambridge: University Press.

Star, C. & Hammer, S. (2008). Teaching generic skills: eroding the higher purpose of universities, or an opportunity for renewal? *Oxford Review of Education*, 34, 2, 237-251.

Sternberg, R. J. & Frensch, P. A. (1992). On being an expert: A cost-benefit analysis. In R. Hoffman (Hrsg.), *The psychology of expertise: Cognitive research and empirical AI*. New York: Springer.

- Sternberg, R. J. (1994). *Thinking and Problem Solving*. London, Kalifornien: Academic Press.
- Sternberg, R. J. (1997). The Concept of Intelligence and Its Role in Lifelong Learning and Success. *American Psychologist*, 52, 10, 1030-1037.
- Sternberg, R. J. & Kaufman, J. C. (1998). Human Abilities. *Annual Review of Psychology*, 49, 479-502.
- Sternberg, R. J. (1999). *The Theory of successful Intelligence*. *Review of General Psychology*, 3, 4, 292-316.
- Sternberg, R. (2003). What Is an “Expert Student”? *Educational Researcher*, 32, 8, 5-9.
- Sternberg, R. J. & Grigorenko, E. L. (2003). *The psychology of abilities, competencies, and expertise*. Cambridge University Press
- Sternberg, R. J. (2003). Biological Intelligence. In R. J. Sternberg & E. L. Grigorenko (Hrsg.), *The psychology of abilities, competencies, and expertise* (240-262). Cambridge University Press.
- Sternberg, R. K. (2005). Intelligence. In K. J. Holyoak & R. G. Morrison (Hrsg.), *The Cambridge Handbook of Thinking and Reasoning* (751-773). Cambridge: University Press.
- Sternberg, R. J., Lubart, T. I., Kaufman, J. C. & Pretz, J. E. (2005). Creativity. In K. J. Holyoak & Morrison, R. (Hrsg.), *Cambridge handbook of thinking and reasoning* (351–369). New York: Cambridge University Press.
- Strike, K. & Posner, G. (1992). A revisionist theory of conceptual change. In R. A. Duschl & R. J. Hamilton (Hrsg.), *Philosophy of science, cognitive psychology, and educational theory and practice* (147–176). Albany: State University Press.
- Strub, Dubach, P. & Oesch, T. (2008). Kompetenzen von Hochschulabsolvent/innen: Anforderungen des Arbeitsmarktes. In Bundesamt für Statistik (Hrsg.), *Schlüsselkompetenzen der Schweizer Hochschulabsolvent/innen. Thematischer Sammelband mit empirischen Ergebnissen der Absolventenstudie* (67-103). Neuchâtel.
- Strube, G. (1998). Modelling motivation and action control in cognitive systems. In U. Schmid, J. Krems & F. Wysotzki (Hrsg.), *Mind modeling* (89-108). Lengerich: Pabst Science Publisher.

- Strube, G. (1990). Neokonnektionismus: Eine neue Basis für die Theorie und Modellierung menschlicher Kognition? *Psychologische Rundschau*, 41, 129-143.
- Sullivan, G. E., (1994). *The Impact of Student Thinking Journals and Generic Problem Solving Software on Problem Solving Performance*. Conference Paper, Annual Meeting of the American Educational Research Association, New Orleans Royer, Cisero und Carlo.
- Suzuki, H., Ohnishi, H. & Shigematsu, K. (1992). *Goal-directed processes in similarity judgment*. Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society, Bloomington, Indiana.
- Süß, H. M., Kersting, M. & Oberauer, K. (1991). Intelligenz und Wissen als Prädiktoren für Leistungen bei computersimulierten komplexen Problemen. *Diagnostica*, 37, 334-352.
- Solso, R. L. (2005). *Kognitive Psychologie*. Heidelberg: Springer.
- Sousa, R. de (2008). Epistemic Feelings. In G. Brun, U. Doğuoğlu & D. Kuenzle (Hrsg.), *Epistemology and emotions* (185-204). Aldershot: Ashgate.
- Souvignier, E., Gold, A. (2006). Wirksamkeit von Lehrmethoden. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (146-166). Heidelberg: Springer.
- Sweller, J. (1994). Cognitive load theory, learning difficulty and instructional design. *Learning and Instruction*, 4, 293-312.
- Thomas, J. (2009). Das Verfahren zum Kompetenznachweis Kultur auf dem Prüfstand. In V. Timmerberg & B. Schorn (Hrsg.), *Neue Wege der Anerkennung von Kompetenzen in der kulturellen Bildung*. München: kopaed.
- Thompson, V. A. (1994). Interpretational factors in conditional reasoning. *Memory & Cognition*, 22, 742-758.
- Thompson, V. A. (2000). The task-specific nature of domain-general reasoning. *Cognition*, 76, 209-268.

- Thorndike, Edward L. (1898). Animal intelligence: An experimental study of the associative processes in animals. *The Psychological Review: Monograph Supplements*, 2, 4, 551-553.
- Tversky, A. & Kahnemann, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- Tversky, A. (1977). Features of Similarity. *Psychological Review*, 84, 327-352.
- Underwood, J., Underwood, G. & Wood, D. (2000). When does gender matter? Interactions during computer-based problem solving. *Learning and Instruction*, 10, 5, 447-462.
- University of Arizona, Higher Education Assessment Tools. Zugriff am 12.01.2014 unter <http://assessment.arizona.edu/tools>
- Urban, D. & Mayerl, J. (o. J.). Mediator-Effekte in der Regressionsanalyse (direkte, indirekte und totale Effekte). Zugriff am 12.01.2014 unter http://www.uni-stuttgart.de/soz/soziologie/regression/Mediator-Effekte_v1-3.pdf
- 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.
- Vaughn, L. (2005). *The power of critical thinking : effective reasoning about ordinary and extraordinary claims*. New York : Oxford University Press. Zugriff am 12.01.2014 unter <http://www.uoup.com/us/companion.websites/0195161424/studentresources/test/?view=usa>
- Van Leer, O. (1987). Upwards and Across: An Essay on Cross-Disciplinary Thinking. 405-412.
- Veenman, M. V. J, Prins, F. J & Elshout, J. J. (2002). Initial inductive learning in a complex computer simulated environment: the role of metacognitive skills and intellectual ability. *Computers in Human Behavior*, 18, 3, 327-341.
- Vendelinski, T. P., Baker, E. L. & Niemi, D. (2008). *Templates and Objects in Authoring Problem-solving Assessment CRESST-Report*. National Center for research on Education, Standards, and Student testing: University of California, Los Angeles.
- Vonken, M. (2001). Von Bildung zur Kompetenz. Die Entwicklung erwachsenenpädagogischer Begriffe oder Rückkehr zur Bildung. *Zeitschrift für Berufs- und Wirtschaftspädagogik*, 4, 501-522.

- Vonken, M. (2005). *Handlung und Kompetenz. Theoretische Perspektiven für die Erwachsenen- und Berufspädagogik*. Wiesbaden: Verlag für Sozialwissenschaften.
- Voss, A., Carstensen, C. H. & Bos, W. (2008). Textgattungen und Verstehensaspekte: Analyse von Leseverständnis aus den Daten der IGLU-Studie. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, R. Valtin & G. Walther (Hrsg.), *IGLU. Vertiefende Analysen zu Leseverständnis, Rahmenbedingungen und Zusatzstudien* (1-36). München: Waxmann.
- Voss, J. F. (1989). Problem solving and the educational process Foundations for a psychology of education. In A. Lesgold, R. Glaser (Hrsg.), *Foundations for a psychology of education* (251-294). Hillsdale, NJ, England: Lawrence Erlbaum Associates.
- Voss, J. F., Greene, T. R., Post, T., & Penner, B. C. (1983). Problem solving skill in the social sciences. In G. Bower (Hrsg.), *The psychology of learning and motivation* (165–213). New York: Academic Press.
- Vye, N. L., Schwartz, D. L., Bransford, J. D., Barron, B. J., Zech, L. & The Cognition and Technology Group at Vanderbilt. (1998). SMART environments that support monitoring, reflection, and revision. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Hrsg.), *Metacognition in educational theory and practice*. Mahwah, NJ: Lawrence Erlbaum.
- Wainer, H. (1992). Understanding graphs and tables. *Educational Researcher*, 21, 1, 14-23.
- Wang, M. C., Haertel, G. D. & Walberg, H. J. (1990). What influences Learning? A Content Analysis of Review Literature. *Journal of Educational research*, 84, 1, 30-38.
- Wang, M. C., Haertel, G. D. & Walberg, H. J. (1993). Toward a Knowledge Base for School Learning. *Review of Educational Research*, 63, 3, 249-294.
- Waldmann, M. R. & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General*, 121, 222-236.
- Wallas, R. A. (1926). *The art of thought*. New York: Hartcourt Brace.
- Wegner, D. M. & Vallacher, R. R. (1986). Action identification. In R. M. Sorrentino & E. T. Higgins (Hrsg.), *Handbook of motivation and cognition: Foundations of social behavior* (550-582). New York: Guilford Press.

- Weidenmann, B. (2001). Lernen mit Medien. In A. Krapp & B. Weidenmann (Hrsg.), *Pädagogische Psychologie* (4., vollständig überarbeitete Auflage Aufl., 423-476). Weinheim: Beltz PVU.
- Weinert, F. E. (2001). Concept of competence: a conceptual clarification. In D. Rychen & L. H. Salganik (Hrsg.), *Defining and selecting key competencies*. Seattle: Hogrefe & Huber.
- Wittmann, W. W., Süß, H. M. & Oberauer, K. (1996). *Determinanten komplexen Problemlösens*. Zugriff am 12.01.2014 unter http://www.psychologie.uni-mannheim.de/psycho2_alt/publi/ps/ber09.pdf
- Wilson, M. (2005). *Constructing Measures: An Item Response Modeling Approach*. Mahwah, NJ: Erlbaum.
- Witmer, A., Storni, M., Schmid, M. (2008). Zur Kompetenzmessung in der Absolventenbefragung: Einleitung und theoretische Konstrukte. In Bundesamt für Statistik (Hrsg.), *Schlüsselkompetenzen der Schweizer Hochschulabsolvent/innen. Thematischer Sammelband mit empirischen Ergebnissen der Absolventenstudie* (19-31). Neuchâtel.
- Wissenschaftsrat (2006). *Empfehlungen zum arbeitsmarkt- und demographiegerechten Ausbau des Hochschulsystems*. Berlin.
- Woodward, R. & Sells, R.B. (1935). An atmosphere effect in formal syllogistic reasoning. *Journal of Experimental Psychology*, 18, 451-460.
- Zimbardo, Ph. G. & Gerrig, R.J. (2004). *Psychologie*. München: Pearson.
- Zimmermann, B.J. & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, 82, 51-59.
- Zlatkin-Troitschanskaia, O. & Kuhn, C. (2010). *Messung akademisch vermittelter Fertigkeiten und Kenntnisse von Studierenden bzw. Hochschulabsolventen – Analyse zum Forschungsstand*. Johannes Gutenberg-Universität Mainz: Arbeitspapiere Wirtschaftspädagogik, 56.

Pilotierung**Tab. 1: Subtest 1 – Problemsituationen analysieren (Vorstudie)**

Items	Schwierigkeit $P_i = n_r/n \cdot 100$	Varianz	Trennschärfe unkorrigiert	ausgewählt
Gesundheit (Ded.)	36,84	,246	,000	x
Familienunternehmen (Ded.)	50,00	,265	,249	x
Berufstätigkeit (Ded.)	68,42	,228	,672	x
Bewerbungskurse (Ded)	68,42	,228	,467	
Wahlkampf	52,63	,263	,456	Analogie nicht aufgenommen
Raucher (F/M)	27,77	,212	,090	x
Gentechnik (F/M)	52,63	,263	,761	emotionaler Aspekt enthalten
Jugendliche Wähler (F/M)	45,00	,261	,415	inhaltlich ungenau
Arbeitslose (F/M)	84,21	,140	,000	x
Politiker (Korr/Kaus)	65,00	,239	,448	
Alkohol (Korr/Kaus)	58,82	,257	,746	x
Neuwagen (Korr/Kaus)	75,00	,197	,393	x
Drogenentwicklung (Korr/Kaus)	65,00	,239	-,170	
Handwerker (Ind.)	50,00	,263	,456	x
Tourismus (Ind)	25,00	,197	,717	Deduktion enthalten
Migration (Ind.)	65,00	,239	,552	x
Dienstleistungssektor (Abd)	75,00	,197	,425	
Migrationsbiographie (Abd)	36,84	,246	,539	Sonderform induktiven Denkens
Nichtraucher (Abd)	90,00	,095	,745	nicht aufgenommen
Reliabilität _{ges} Cronbachs Alpha .74				

Tab. 2: Subtest 1 – Problemsituationen analysieren (Hauptstudie) – Einzelitemebene

Items	Schwierigkeit $P_i = n_r/n \cdot 100$	Varianz	Trennschärfe unkorrigiert
Gesundheit (Ded.)	17,62	,146	,093
Berufstätigkeit (Ded.)	28,16	,203	,298
Familienunternehmen (Ded.)	58,30	,244	,107
Handwerker (Ind.)	63,25	,233	,623
Migration (Ind.)	73,33	,196	,568
Alkohol (Korr/Kaus)	60,42	,240	,518
Neuwagen (Korr/Kaus)	77,19	,177	,432
Raucher (F/M)	26,57	,196	,387
Arbeitslose (F/M)	75,8	,184	,321

Reliabilität_{ges} Cronbachs Alpha .29

Tab. 3: Subtest 1 – Problemsituationen analysieren (Hauptstudie) – Itemsetebene

Items	Schwierigkeit $P_i = n_r/n \cdot 100$	Varianz	Trennschärfe unkorrigiert
Fakten vs. Meinung	22,37	,470	,488
Deduktion	3,49 (alle Items richtig)		
	69,58 (≥ 2 Items richtig)	,733	,352
Korrelation vs. Kausalität	50,69	,538	,629
Induktion	57,69	,710	,651

Reliabilität_{ges} Cronbachs Alpha .13

Reliabilität_{exkl} Cronbachs Alpha .51 (exkl. Itemset Deduktion)

Tab. 4: Subtest 2 – Ziele bewerten und evaluieren (Vorstudie) – Einzelitemebene

Items	Schwierigkeit* $P_i = n_r/n \cdot 100$	Varianz	Trennschärfe unkorrigiert	aus- gewählt
Lohnnebenkosten	25,00	,526	,565	x
Verwaltungsapparat	45,00	,358	,559	x
Marode Gebäude	45,00	,358	,664	
Beispielingen 42	50,00	,726	,621	

Reliabilität_{ges} Cronbachs Alpha .39

*n_r= nur diejenigen, die das Item vollständig gelöst haben, d.h. beide Kriterien korrekt evaluiert haben

Tab. 5: Subtest 2 – Ziele bewerten und evaluieren (Hauptstudie) – Einzelitemebene

Items	Schwierigkeit* $P_i = n_r/n \cdot 100$	Varianz	Trennschärfe unkorrigiert
Lohnnebenkosten	48,65	,518	-
Verwaltungsapparat	49,44	,288	-

*n_r= nur diejenigen, die das Item vollständig gelöst haben, d.h. beide Kriterien korrekt evaluiert haben.

Tab. 6: Subtest 3 – Handlungen planen (Vorstudie) – Einzelitemebene

Items	Schwierigkeit* $P_i = n_r/n \cdot 100$	Varianz	Trennschärfe unkorrigiert	aus- gewählt
Arbeitslosigkeit senken	13,33	2,69	.835	x
Getränkehersteller	12,5	3,71	.757	x
Flussbett	33,33	3,25	.747	
Bundesgartenschau	12,5	1,58	.579	

Reliabilität_{ges} Cronbachs Alpha .71

*n_r=nur diejenigen die das Item vollständig und korrekt gelöst haben

Tab. 7: Subtest 3 – Handlungen planen (Hauptstudie) – Einzelitemebene

Items Subtest 3	Schwierigkeit $P_i = n_r/n \times 100$	Varianz	Trennschärfe unkorrigiert
Arbeitslosigkeit senken	16,90	1,19	-
Getränkehersteller	9,21	0,95	-

Tab. 8: Subtest 3 – Handlungen planen (Hauptstudie) – Einzelitemebene

Items	Schwierigkeit* $P_i = n_r/n \cdot 100$	Varianz	Trennschärfe unkorrigiert
Arbeitslosigkeit senken	3,52	1,79	-
Getränkehersteller	8,86	1,77	-

* n_r =richtige Platzierung der Maßnahmen sowie korrektes Erkennen beider Distraktoren

Tab. 9: Subtest 3 – Handlungen planen (Hauptstudie) – Distraktoren Item *Arbeitslosigkeit*

Erkennen der Distraktoren		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	keiner	78	26,9	27,5	27,5
	beide	59	20,3	20,8	48,2
	distaler Distraktor	51	17,6	18,0	66,2
	proximaler Distraktor	96	33,1	33,8	100,0
	Gesamt	284	97,9	100,0	
Fehlend		6	2,1		
Gesamt		290	100,0		

Tab. 10: Subtest 3 – Handlungen planen (Hauptstudie) – Distraktoren Item *Getränkehersteller*

Erkennen der Distraktoren		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	keiner	57	19,7	20,2	20,2
	beide	105	36,2	37,2	57,4
	distaler Distraktor	35	12,1	12,4	69,9
	proximaler Distraktor	85	29,3	30,1	100,0
	Gesamt	282	97,2	100,0	
Fehlend		8	2,8		
Gesamt		290	100,0		

Tab. 11. Analyse fehlender Werte für die Gesamtstichprobe (Hauptstichprobe) – Ebene der Itemsets

Itemset	N	Mittelwert	Standardabweichung	Fehlend		Anzahl der Extremwerte ^a		
				Anzahl	Prozent	Niedrig	Hoch	
Subtest 1								
Fakten vs. Meinung	28 6	3,00	1,86	0	,0	42	62	
Korrelation vs. Kausalität	28 6	4,17	1,79	0	,0	0	0	
Deduktion	28 6	3,24	2,01	0	,0	0	0	
Induktion	28 6	4,10	1,90	0	,0	0	0	
Subtest 2								
Zielevaluation	27 1	2,75	,98	15	5,2	3	0	
Subtest 3								
Handlungsplanung	28 3	3,78	2,16	3	1,0	0	5	

a. Anzahl der Fälle außerhalb des Bereichs ($Q1 - 1,5 \cdot IQR$, $Q3 + 1,5 \cdot IQR$).

Tab. 12: Analyse fehlender Werte (Hauptstichprobe) – Übersicht auf Einzelitemebene

	N	Mittelwert	Standardabweichung	Fehlend		Anzahl der Extremwerte ^{b,a}	
				Anzahl	Prozent	Niedrig	Hoch
Arbeitslose	282	2,67	,917	9	3,1	.	.
Neuwagenkäufe	285	2,76	,803	6	2,1	.	.
Raucher	286	2,56	,915	5	1,7	0	0
Berufstätigkeit	284	2,69	,999	7	2,4	0	0
Handwerker	283	2,68	,977	8	2,7	0	0
Migrationsbiographie	285	2,81	,793	6	2,1	0	0
Familienunternehmen	283	2,24	1,114	8	2,7	0	0
Gesundheits-	278	3,02	,748	13	4,5	.	.

kosten							
Alkoholpreise	283	2,60	,945	8	2,7	0	0
Verwaltungs- apparat	271	3,41	,703	20	6,9	7	0
Lohnneben- kosten	265	1,98	1,148	26	8,9	0	0
Arbeitslosigkeit	284	,94	1,091	7	2,4	0	48
Getränke- hersteller	282	,72	,952	9	3,1	0	26

a. . zeigt an, daß der IQR (Bereich zwischen den Quartilen) gleich Null ist.

b. Anzahl der Fälle außerhalb des Bereichs ($Q1 - 1,5 \cdot IQR$, $Q3 + 1,5 \cdot IQR$).

Tab. 13: Analyse fehlender Werte – Semester 6

	N	Mittel- wert	Standard- abweichung	Fehlend		Anzahl der Extremwerte ^{b,a}	
				Anzahl	Prozent	Niedrig	Hoch
Arbeitslose	97	2,75	,902	2	2,0	.	.
Neuwagenkäufe	98	2,86	,718	1	1,0	.	.
Raucher	99	2,57	,917	0	,0	0	0
Berufstätigkeit	99	2,72	,915	0	,0	0	0
Handwerker	97	2,74	1,003	2	2,0	0	0
Migrationsbiogra- phie	99	2,69	,737	0	,0	0	0
Familienunter- nehmen	99	2,38	1,122	0	,0	0	0
Gesundheitskosten	95	3,11	,722	4	4,0	4	0
Alkoholpreise	98	2,87	,808	1	1,0	0	0
Verwaltungs- apparat	95	3,44	,664	4	4,0	2	0
Lohnnebenkosten	94	2,07	1,138	5	5,1	0	0
Arbeitslosigkeit	99	1,07	1,062	0	,0	0	0
Getränkehersteller	99	,76	,980	0	,0	0	9

a. . zeigt an, daß der IQR (Bereich zwischen den Quartilen) gleich Null ist.

b. Anzahl der Fälle außerhalb des Bereichs ($Q1 - 1,5 \cdot IQR$, $Q3 + 1,5 \cdot IQR$).

Tab. 14: Analyse fehlender Werte – Semester 3

	N	Mittelwert	Standardabweichung	Fehlend		Anzahl der Extremwerte ^b	
				Anzahl	Prozent	Niedrig	Hoch
Arbeitslose	69	2,54	,964	1	1,4	0	0
Neuwagenkäufe	70	2,77	,726	0	,0		
Raucher	70	2,47	,944	0	,0	0	0
Berufstätigkeit	69	2,65	1,027	1	1,4	0	0
Handwerker	70	2,60	,907	0	,0	0	0
Migrationsbiographie	70	2,91	,737	0	,0		
Familienunternehmen	68	2,26	1,192	2	2,9	0	0
Gesundheitskosten	69	3,06	,802	1	1,4	5	0
Alkoholpreise	69	2,49	,918	1	1,4	0	0
Verwaltungsapparat	66	3,42	,725	4	5,7	2	0
Lohnnebenkosten	62	2,21	1,161	8	11,4	0	0
Arbeitslosigkeit	70	,91	1,073	0	,0	0	10
Getränkehersteller	69	,71	,972	1	1,4	0	7

a. . zeigt an, daß der IQR (Bereich zwischen den Quartilen) gleich Null ist.

b. Anzahl der Fälle außerhalb des Bereichs ($Q1 - 1,5 \cdot IQR$, $Q3 + 1,5 \cdot IQR$).

Tab. 15: Analyse fehlender Werte – Semester 1

	N	Mittelwert	Standardabweichung	Fehlend		Anzahl der Extremwerte ^{b,a}	
				Anzahl	Prozent	Niedrig	Hoch
Arbeitslose	113	2,67	,911	1	,9		
Neuwagenkäufe	114	2,68	,917	0	,0	0	0
Raucher	114	2,63	,885	0	,0		
Berufstätigkeit	113	2,69	1,070	1	,9	0	0
Handwerker	113	2,65	,998	1	,9	0	0
Migrationsbiographie	113	2,84	,872	1	,9	0	0
Familienunternehmen	113	2,12	1,059	1	,9	0	0
Gesundheitskosten	111	2,90	,726	3	2,6		
Alkoholpreise	113	2,44	1,035	1	,9	0	0
Verwaltungsapparat	107	3,35	,728	7	6,1	3	0
Lohnnebenkosten	106	1,70	1,071	8	7,0	0	0
Arbeitslosigkeit	112	,85	1,117	2	1,8	0	20
Getränkehersteller	111	,71	,928	3	2,6	0	10

a. . zeigt an, daß der IQR (Bereich zwischen den Quartilen) gleich Null ist.

b. Anzahl der Fälle außerhalb des Bereichs ($Q1 - 1,5 \cdot IQR$, $Q3 + 1,5 \cdot IQR$).

Tab. 16: Analyse fehlender Werte – Geisteswissenschaften

	N	Mittelwert	Standard- abweichung	Fehlend		Anzahl der Extremwerte ^a	
				Anzahl	Prozent	Niedrig	Hoch
Arbeitslose	126	2,56	,976	4	3,1	0	0
Neuwagenkäufe	130	2,73	,824	0	,0	0	0
Raucher	130	2,57	,940	0	,0	0	0
Berufstätigkeit	128	2,59	1,061	2	1,5	0	0
Handwerker	127	2,59	1,072	3	2,3	0	0
Migrations- biographie	130	2,75	,854	0	,0	0	0
Familienunter- nehmen	127	2,18	1,094	3	2,3	0	0
Gesundheits- kosten	124	3,00	,776	6	4,6	6	0
Alkoholpreise	128	2,52	1,019	2	1,5	0	0
Verwaltungs- apparat	124	3,39	,783	6	4,6	5	0
Lohnnebenkosten	121	1,94	1,199	9	6,9	0	0
Arbeitslosigkeit	129	,86	1,051	1	,8	0	19
Getränke- hersteller	127	,59	,858	3	2,3	0	8

a. Anzahl der Fälle außerhalb des Bereichs ($Q1 - 1,5 \cdot IQR$, $Q3 + 1,5 \cdot IQR$).

Tab. 17: Analyse fehlender Werte – Sozialwissenschaften

	N	Mittelwert	Standard- abweichung	Fehlend		Anzahl der Extremwerte ^{b,a}	
				Anzahl	Prozent	Niedrig	Hoch
Arbeitslose	156	2,76	,860	0	,0	.	.
Neuwagenkäufe	155	2,79	,787	1	,6	.	.
Raucher	156	2,55	,897	0	,0	0	0
Berufstätigkeit	156	2,78	,941	0	,0	.	.
Handwerker	156	2,76	,890	0	,0	0	0
Migrations- biographie	155	2,85	,737	1	,6	.	.
Familien- unternehmen	156	2,29	1,131	0	,0	0	0
Gesundheitskosten	154	3,03	,727	2	1,3	.	.
Alkoholpreise	155	2,68	,875	1	,6	0	0
Verwaltungs- apparat	147	3,42	,629	9	5,8	2	0
Lohnnebenkosten	144	2,01	1,106	12	7,7	0	0
Arbeitslosigkeit	155	1,01	1,122	1	,6	0	0
Getränkehersteller	155	,83	1,012	1	,6	0	18

a. . zeigt an, daß der IQR (Bereich zwischen den Quartilen) gleich Null ist.

b. Anzahl der Fälle außerhalb des Bereichs ($Q1 - 1,5 \cdot IQR$, $Q3 + 1,5 \cdot IQR$).

Tab. 18: Kennwerte Gesamttest

Subtests (z-transformiert)	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Quadrierte multiple Korrelation	Cronbachs Alpha, wenn Item weggelassen
Subtest Situationsanalyse	2,481	0,98	0,11	,388
Subtest Zielevaulation	2,207	,196	,059	,188
Subtest Handlungsplanung	2,100	2,38	0,66	0,95
Reliabilität _{ges} Cronbachs Alpha .31				

Tab. 19: Studieninteresse – Items, Dimension und Polung

Ich bin mir sicher, das Fachstudium gewählt zu haben, welches meinen persönlichen Neigungen entspricht. (+)* *intrinsischer Charakter***

Ich rede lieber über meine Hobbies als über mein Studienfach. (-) *gefühlbezogene Valenzen*

Wenn ich in einer Bibliothek oder in einem Buchladen bin schmökere ich gerne in Zeitschriften oder Büchern, die Themen aus meinem Studienfach ansprechen. (+) *gefühlbezogene Valenzen*

Ich habe mein jetziges Fachstudium vor allem wegen der interessanten Studieninhalte gewählt. (+) *intrinsischer Charakter*

Wenn ich ehrlich sein soll, ist mir mein Studienfach manchmal eher gleichgültig. (-) *persönliche wertbezogene Valenzen*

Im Vergleich zu anderen mir sehr wichtigen Dingen (z. B. Hobbies, soziale Beziehungen) messe ich meinem Fachstudium eher eine geringe Bedeutung bei. (-) *persönliche wertbezogene Valenzen*

In Klammern angegeben ist die Polung des Items (*), kursiv die Dimension des Items (**)
Skala nach Schiefele, Krapp & Winterler (1993). Der "Fragebogen zum Studieninteresse" (FSI).

Tab. 20: Selbstwirksamkeit – Items und Polung

Ich kann auch die schwierigen Aufgaben in meinem Studienfach lösen, wenn ich mich anstrengende. (+)*

Es fällt mir leicht, neuen Unterrichtsstoff zu verstehen. (+)

Wenn ich im Rahmen meines Studienfachs eine schwierige Aufgabe lösen soll, glaube ich, daß ich das schaffen werde. (+)

Selbst wenn ich mal längere Zeit krank sein sollte, kann ich immer noch gute Leistungen erzielen. (+)

Ich bin mir sicher, dass ich auch dann noch meine gewünschten Leistungen erreichen kann, wenn ich mal eine schlechte Note bekommen habe. (+)

Wenn der Dozent / die Dozentin das Tempo noch mehr anzieht, werde ich die geforderten Leistungen kaum noch schaffen können. (-)

Auch wenn der Dozent / die Dozentin an meinen Fähigkeiten zweifelt, bin ich mir sicher, dass ich gute Leistungen erzielen kann. (+)

In Klammern angegeben ist die Polung des Items (*)
Skala nach Jerusalem & Satow (1999).

Tab. 21: Relevanz fächerübergreifenden Problemlösens – Items und Polung

In meinem angestrebten Beruf werde ich sehr sicher auch mit Problemen konfrontiert, die eher nichts mit meinem jetzigen Fachstudium zu tun haben. (+)*

Wenn ich in meinem Beruf erfolgreich sein möchte, so muss ich auch sehr gut mit solchen Problemstellungen umgehen können. (+)

In meinem Beruf muss ich diese Art von Problemstellungen sicher nicht alleine bewältigen, da ich bspw. immer gemeinsam im Team mit anderen Kollegen an Lösungen arbeiten kann. (-)

Ich bin sicherlich in meinem Beruf auch erfolgreich, wenn ich ausschließlich Experte in meinem studierten Fachbereich bin. (-)

In Klammern angegeben ist die Polung des Items (*)

Tab. 22: Transfer – Items und Polung

Wenn ich vor Problemen stehe, die eher nichts mit meinem Fachstudium zu tun haben, helfen mir die **Herangehensweisen aus meinem Fachstudium** dabei, dass ich diese Probleme zuerst einmal gut analysieren und verstehen kann. (Bspw. weiß ich aufgrund meines Fachstudiums, wie ich diese Probleme erkennen und analysieren könnte) (+)*

Die **Herangehensweisen**, die mir **aus meinem Fachstudium** bekannt sind, helfen mir dabei, dass ich mir für diese Art von Problemen sehr gut Lösungsmöglichkeiten überlegen kann. (Bspw. weiß ich aufgrund meines Fachstudiums, wie ich diese Probleme grundsätzlich angehen könnte) (+)

Die **Arbeitsmethoden zur Lösung von fachlichen Problemen**, die mir aus meinem Fachstudium bekannt sind, helfen mir dabei, mit dieser Art von Problemstellungen sehr gut umgehen zu können. (Bspw. weiß ich deshalb, wie ich diese Probleme strategisch bearbeiten könnte) (+)

Das **Wissen**, das ich mir im Rahmen **meines Fachstudiums** erarbeitet habe, hilft mir dabei, dass ich mit diesen Problemen besser umgehen kann. (Bspw. kann ich diese Probleme und Fragestellungen deshalb besser verstehen, da mir mein Fachwissen dazu hilfreiche Ansatzpunkte gibt) (+)

In Klammern angegeben ist die Polung des Items (*)

Tab. 23: Epistemologische Beliefs: Quelle des Wissens – Items und Polung

Was man in Büchern meiner Fachwissenschaft liest, ist sicher wahr. (-)*
Was der Dozent im Unterricht sagt, ist sicher wahr. (-)
Was Wissenschaftler meines Studienfaches herausfinden, muss man glauben. (-)
Was in Büchern meiner Fachwissenschaft steht, muss man glauben. (-)
Nur Wissenschaftler meines Faches wissen genau, was in meinem Studienfach wahr ist. (-)
In Klammern angegeben ist die Polung des Items (*) Adaptierte Skala nach Urhahne & Hopf (2004).

Tab. 24: Epistemologische Beliefs: Sicherheit des Wissens – Items und Polung

Der wichtigste Teil der Fachwissenschaft, die ich studiere, ist die Suche nach den einzig richtigen Lösungen. (-)*
Es gibt nur die eine Lösung, wenn Wissenschaftler meiner Fachwissenschaft einmal das Ergebnis einer Untersuchung gefunden haben. (-)
In meiner Fachwissenschaft ist beinahe alles bekannt; es gibt nicht mehr viel, was man herausfinden könnte. (-)
Alle Fragen in meiner Fachwissenschaften haben genau eine Lösung. (-)
Die Fachwissenschaftler meines Studienfaches stimmen immer darin überein, was in meinem Fach wahr ist. (-)
Das Wissen in der Fachwissenschaft, die ich studiere ist für alle Zeit wahr. (-)
In Klammern angegeben ist die Polung des Items (*) Adaptierte Skala nach Urhahne & Hopf (2004).

Tab. 25: Elterlicher Bildungshintergrund – Hauptstudie

		Höchster Bildungsabschluss des Vaters			
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Hauptschulabschluss	55	19,0	21,1	21,1
	Realschulabschluss	56	19,3	21,5	42,5
	Abitur	39	13,4	14,9	57,5
	Studienabschluss	88	30,3	33,7	91,2
	Promotion	23	7,9	8,8	100,0
	Gesamt	261	90,0	100,0	
Fehlend	System	29	10,0		
Gesamt		290	100,0		

		Höchster Bildungsabschluss der Mutter			
		Häufigkeit	Prozent	Gültige Prozente	Kumulierte Prozente
Gültig	Hauptschulabschluss	50	17,2	19,3	19,3
	Realschulabschluss	102	35,2	39,4	58,7
	Abitur	28	9,7	10,8	69,5
	Studienabschluss	72	24,8	27,8	97,3
	Promotion	7	2,4	2,7	100,0
	Gesamt	259	89,3	100,0	
Fehlend	System	31	10,7		
Gesamt		290	100,0		

Sehr geehrte Teilnehmerinnen und Teilnehmer,

der Ihnen vorliegende Fragebogen ist Teil einer Promotionsstudie. Das Ziel der Studie ist, die Lehre an Universitäten zu verbessern, um Sie möglichst optimal auf Ihre künftigen beruflichen Anforderungen vorzubereiten.

In diesem Teil der Studie geht es darum, zu erfahren, **für wie wichtig** Sie als Studierende die Fähigkeit zum fachübergreifenden Problemlösen halten. Fachübergreifendes Problemlösen bedeutet, auch Probleme lösen zu können, die nicht explizit Teil Ihres Studiums waren (bspw. mit komplexen Informationen umgehen und diese strukturieren können). Da diese Art von Problemen ohne Fachwissen gelöst werden kann und nicht einem bestimmten Fachbereich zuzuordnen ist, werden sie als fachübergreifende Probleme bezeichnet.

Daneben ist es im Rahmen der Studie auch sehr wichtig, zu erfahren, wie Sie Ihr Studium insgesamt einschätzen.

Die Beantwortung dieses Teils der Studie wird insgesamt **ca. 10 Minuten** Zeit in Anspruch nehmen. Natürlich sind alle Angaben freiwillig.

Da alle Angaben im Rahmen der Studie anonym sind und keine Rückschlüsse auf Ihre Person gezogen werden können, ist es **sehr wichtig**, dass Sie sich ein **Kennwort nach dem nachfolgenden Schema** überlegen. Es dient dazu, die beiden Teile der Studie wieder zuzuordnen.

Ihr Kennwort:

Die ersten zwei Buchstaben des Vornamens Ihrer Mutter =>Beispiel: Monika = Mo	Die ersten zwei Buchstaben des Vornamens Ihres Vaters =>Beispiel: Werner = We	Die ersten zwei Ziffern Ihres eigenen Geburtstages (nur der Tag) => Beispiel: 03.12.1984 = 03	Die zwei Buchstaben Ihres eigenen Geburtsortes =>Beispiel: Dresden = Dr

	Trifft gar nicht zu	Trifft kaum zu	Trifft eher zu	Trifft genau zu
In meinem angestrebten Beruf werde ich sehr sicher auch mit Problemen konfrontiert, die eher nichts mit meinem jetzigen Fachstudium zu tun haben.				
Wenn ich in meinem Beruf erfolgreich sein möchte, so muss ich auch sehr gut mit solchen Problemstellungen umgehen können.				
In meinem Beruf muss ich diese Art von Problemstellungen sicher nicht alleine bewältigen, da ich bspw. immer gemeinsam im Team mit anderen Kollegen an Lösungen arbeiten kann.				
Ich bin sicherlich in meinem Beruf auch erfolgreich, wenn ich ausschließlich Experte in meinem studierten Fachbereich bin.				
Wenn ich vor Problemen stehe, die eher nichts mit meinem Fachstudium zu tun haben, helfen mir die Herangehensweisen aus meinem Fachstudium dabei, dass ich diese Probleme zuerst einmal gut analysieren und verstehen kann. (Bspw. weiß ich aufgrund meines Fachstudiums, wie ich diese Probleme erkennen und analysieren könnte)				
Die Herangehensweisen , die mir aus meinem Fachstudium bekannt sind, helfen mir dabei, dass ich mir für diese Art von Problemen sehr gut Lösungsmöglichkeiten überlegen kann. (Bspw. weiß ich aufgrund meines Fachstudiums, wie ich diese Probleme grundsätzlich angehen könnte)				
Die Arbeitsmethoden zur Lösung von fachlichen Problemen , die mir aus meinem Fachstudium bekannt sind, helfen mir dabei, mit dieser Art von Problemstellungen sehr gut umgehen zu können. (Bspw. weiß ich deshalb, wie ich diese Probleme strategisch bearbeiten könnte)				
Das Wissen , das ich mir im Rahmen meines Fachstudiums erarbeitet habe, hilft mir dabei, dass ich mit diesen Problemen besser umgehen kann. (Bspw. kann ich diese Probleme und Fragestellungen deshalb besser verstehen, da mir mein Fachwissen dazu hilfreiche Ansatzpunkte gibt)				
Die von mir im Fachstudium gelernten Herangehensweisen zur Lösung von fachlichen Problemen bieten sehr gute Ansatzpunkte zur Lösung von Problemstellungen, die eher nichts mit meinem Fachstudium zu tun haben.				
Mein Fachstudium soll mich in erster Linie sehr gut studienfachspezifisch ausbilden, außerfachliche Problemstellungen sollen eher nicht im Vordergrund stehen (Bsp. Die Uni legt eher den theoretischen Grundstock für die spätere Praxis).				
Mein Fachstudium kann mich in erster Linie nur sehr gut studienfachspezifisch ausbilden, außerfachliche Problemstellungen können eher nicht im Rahmen meines Fachstudiums vermittelt werden (Bsp. Uni ist Uni, Praxis ist Praxis; beide können nur jeweils das eine leisten).				
Mit Problemstellungen umgehen zu können, die eigentlich nichts mit meinem Fachstudium zu tun haben, lerne ich sehr viel besser in der beruflichen Praxis (Praktika, Trainee, spätere berufliche Tätigkeit).				
Mit Problemstellungen umgehen zu können, die eher nichts mit meinem Studium zu tun haben, lerne ich sehr viel besser in praxisbezogenen Veranstaltungen (Bspw. von Praktikern geleitete Workshops), die im Rahmen meines Fachstudiums angeboten werden.				

Sie finden im Folgenden eine Reihe von Aussagen, die sich ebenfalls wieder auf Ihr Studium und ihr Studienfach beziehen. Es geht auch hier um Ihre **persönlichen Einschätzungen und Gefühle**. Gehen Sie bitte die Aussagen der Reihe nach durch. Kreuzen Sie bei jeder Aussage das Kästchen an („trifft gar nicht zu“ bis „trifft genau zu“), das am ehesten zutrifft.

	Trifft gar nicht zu	Trifft kaum zu	Trifft eher zu	Trifft genau zu
Ich bin mir sicher, das Studienfach gewählt zu haben, welches meinen persönlichen Neigungen entspricht.				
Im Vergleich zu anderen mir sehr wichtigen Dingen (z. B. Hobbies, soziale Beziehungen) messe ich meinem Fachstudium eher eine geringe Bedeutung bei.				
Wenn der Dozent / die Dozentin das Tempo noch mehr anzieht, werde ich die geforderten Leistungen kaum noch schaffen können.				
Wenn ich in einer Bibliothek oder in einem Buchladen bin schmökere ich gerne in Zeitschriften oder Büchern, die Themen aus meinem Studienfach ansprechen.				
Ich kann auch die schwierigen Aufgaben in meinem Studienfach lösen, wenn ich mich anstrenge.				
Das Wissen in der Fachwissenschaft, die ich studiere ist für alle Zeit wahr.				
Wenn ich ehrlich sein soll, ist mir mein Studienfach manchmal eher gleichgültig.				
Der wichtigste Teil der Fachwissenschaft, die ich studiere, ist die Suche nach den einzig richtigen Lösungen.				
Wenn ich im Rahmen meines Studienfachs eine schwierige Aufgabe lösen soll, glaube ich, dass ich das schaffen werde.				
Es gibt nur die eine Lösung, wenn Wissenschaftler meiner Fachwissenschaft einmal das Ergebnis einer Untersuchung gefunden haben.				
Was man in Büchern meiner Fachwissenschaft liest, ist sicher wahr.				
Was der Dozent im Unterricht sagt, ist sicher wahr.				
Ich habe mein jetziges Fachstudium vor allem wegen der interessanten Studieninhalte gewählt.				
Was in Büchern meiner Fachwissenschaft steht, muss man glauben.				
Ich bin mir sicher, dass ich auch dann noch meine gewünschten Leistungen erreichen kann, wenn ich mal eine schlechte Note bekommen habe.				
Was Wissenschaftler meines Studienfaches herausfinden, muss man glauben.				
Selbst wenn ich mal längere Zeit krank sein sollte, kann ich immer noch gute Leistungen erzielen.				
In meiner Fachwissenschaft ist beinahe alles bekannt; es gibt nicht mehr viel, was man herausfinden könnte.				

	Trifft gar nicht zu	Trifft kaum zu	Trifft eher zu	Trifft genau zu
Alle Fragen in meiner Fachwissenschaften haben genau eine Lösung.				
Auch wenn der Dozent / die Dozentin an meinen Fähigkeiten zweifelt, bin ich mir sicher, dass ich gute Leistungen erzielen kann.				
Nur Wissenschaftler meines Faches wissen genau, was in meinem Studienfach wahr ist.				
Ich rede lieber über meine Hobbies als über mein Studienfach.				
Die Fachwissenschaftler meines Studienfaches stimmen immer darin überein, was in meinem Fach wahr ist.				
Es fällt mir leicht, neuen Stoff zu verstehen.				

Demographische Daten:

Weiblich

Männlich

Alter (Jahre)

Abiturnote (Bsp.: „2,9“)

Höchster Bildungsabschluss der Eltern

- Vater:
- Hauptschulabschluss
 - Realschulabschluss
 - Abitur
 - Studienabschluss
 - Promotion

- Mutter:
- Hauptschulabschluss
 - Realschulabschluss
 - Abitur
 - Studienabschluss
 - Promotion

Dieses Studium ist mein Erststudium:

Ja

Nein

Ich habe bereits eine Berufsausbildung abgeschlossen.

Ja

Nein

Ich habe bereits im Rahmen einer **regulären Anstellung** Berufserfahrung reguläre berufliche Tätigkeit gesammelt:

ca.Monate

(gemeint ist **nicht eine ggf. absolvierte Ausbildungszeit, sondern bspw. die Zeit nach der Ausbildung**)

In meinem Fachstudium habe ich im
Rahmen von **Praktika** bereits
Berufserfahrung gesammelt: ca.Wochen

Auf einer Skala von 1: „keine berufliche
Erfahrung“ - 4 „beruflich sehr erfahren“
schätze ich meine berufliche Erfahrung
ein als:
.....
(1: „keine berufliche Erfahrung“ 2 – 3- 4: „beruflich sehr erfahren“)

Ich danke Ihnen sehr für Ihre Unterstützung!

Tabelle 26: Normalverteilung und Varianzhomogenität für die Gesamtstichprobe

Itemset	M	SD	s	Min	Max	95% Konfidenz- intervall	Schiefe (SE)	Kurtosis (SE)	Signifikanz		Signifikanz
									Kolmogorov -Smirnov	Shapiro Wilk	Levene
Induktion	1,36	,82	,67	,00	2,00	1,26-1,45	-,74 (,14)	-1,11 (,29)	,00	,00	,11
Deduktion	1,03	,81	,66	,00	3,00	,93-1,12	,35 (,14)	-,55 (,29)	,00	,00	,86
Korrelation und Kausalität	1,37	,72	,51	,00	2,00	1,28-1,45	-,68 (,14)	-,80 (,29)	,00	,00	,90
Fakten und Meinung	1,01	,66	,44	,00	2,00	,94-1,09	-,02 (,14)	-,68 (,29)	,00	,00	,03
Situations- analyse	4,77	1,56	2,45	,00	8,00	4,58-4,95	-,40 (,14)	,01 (,29)	,00	,00	,12
Ziel- evaluation	2,82	,05	,83	,00	4,00	2,72-2,93	-,55 (,14)	-,16 (,29)	,00	,00	,21
Handlungs- planung	3,75	2,17	4,71	,00	10,00	3,50-4,00	,69 (,14)	,42 (,29)	,00	,00	,48
Gesamttest	11,34	3,13	9,79	2,00	19,00	10,98-11,70	,06 (,14)	,01 (,29)	,00	,028	,06
N=286											

Tabelle 27: Nichtparametrische Tests für unabhängige Stichproben

Itemset	Kolmogorov-Smirnov-Z*	Signifikanz	Kolmogorov-Smirnov-Z**	Signifikanz
Induktion	,95	,33	,13	1,00
Deduktion	,76	,62	1,33	,06
Korrelation und Kausalität	1,06	,21	1,79	1,00
Fakten und Meinung	,92	,37	,32	,00
Situationsanalyse	,65	,65	,58	,58
Zielevaluation	,54	,93	,55	,55
Handlungsplanung	,90	,40	,45	,45
Gesamttest	1,06	,21	,97	,97

*nichtparametrischer Test für unabhängige Stichproben. Nullhypothese: Beide Stichproben (Geisteswiss. u. Sozialwiss.) entstammen derselben Grundgesamtheit (Geisteswiss. N=130, Sozialwiss. N=156).

**nichtparametrischer Test für unabhängige Stichproben. Nullhypothese: Beide Stichproben (1. und 6. Semester) entstammen derselben Grundgesamtheit (1. Semester: N=114, 6. Semester N=102).

Fragenkomplex I:**Exploratorische Faktorenanalyse****Tabelle 28: Einfaktorielle Hauptkomponentenanalyse (Subtestebene)**

Einfaktorielle Struktur	h^2	Faktor 1 ($\delta^2 = 42,7\%$)
Situationsanalyse	.162	.402
Zielevaluation	.528	.726
Handlungsplanung	.591	.769

N= 286; Hauptkomponentenanalyse; Varimaxrotation; KMO = .518; Bartlett Test auf Sphärizität .000;
Gesamtvarianz $\delta^2_{\text{Gesamt}} = 43\%$; Eigenwerte $\lambda = 1.28$;

Tabelle 29: Einfaktorielle Hauptachsenanalyse (Subtestebene)

Einfaktorielle Struktur	h^2	Faktor 1 ($\delta^2 = 42,7\%$)
Situationsanalyse	.024	.154
Zielevaluation	.150	.388
Handlungsplanung	.378	.615

N= 286; Hauptachsenanalyse; Varimaxrotation; KMO = .518; Bartlett Test auf Sphärizität .000;
Gesamtvarianz $\delta^2_{\text{Gesamt}} = 43\%$; Eigenwerte $\lambda = 1.28$;

Tabelle 30: Zweifaktorielle Hauptachsenanalyse (Subtestebene)

Zweifaktorielle Struktur	h^2	Faktor 1 ($\delta^2= 42,7\%$)	Faktor 2 ($\delta^2= 32,2\%$)
Situationsanalyse	.082	.126	.286
Zielevaluation	.239	.486	.160
Handlungsplanung	.282	.510	.351

N= 286; Hauptachsenanalyse; Oblimin-direkt; KMO = .518; Bartlett Test auf Sphärizität .000;

Gesamtvarianz $\delta^2_{\text{Gesamt}} = 75\%$ *; Eigenwerte $\lambda = 1.28; .967$;

* Die Angabe einer Gesamtvarianz lässt aufgrund der Berücksichtigung der Interkorrelation der Faktoren keinen aussagekräftigen Schluss zu.

Tabelle 31: Zweifaktorielle Hauptkomponentenanalyse (Itemsetebene)

Zweifaktorielle Struktur	h^2	Faktor 1 ($\delta^2= 27,3\%$)	Faktor 2 ($\delta^2= 21,2\%$)
Induktion	.549	.740	
Korrelation	.616	.775	
Deduktion	.302	-.549	
Fakten und Meinung	.305	.332	.442
Zielevaluation	.555		.727
Handlungsplanung	.588		.765

N= 286; Hauptkomponentenanalyse; Varimaxrotation; KMO = .565; Bartlett Test auf Sphärizität .000;

Gesamtvarianz $\delta^2_{\text{Gesamt}} = 49\%$; Eigenwerte $\lambda = 1.64; 1.28$;

Tabelle 32: Dreifaktorielle Hauptkomponentenanalyse (Itemsetebene)

Dreifaktorielle Struktur	h^2	Faktor 1 ($\delta^2= 27,3\%$)	Faktor 2 ($\delta^2= 21,2\%$)	Faktor 3 ($\delta^2= 15,3\%$)
Induktion	.746	.863		
Korrelation	.672	.786		
Fakten und Meinung	.629			.749
Deduktion	.552			-.680
Zielevaluation	.646		.790	
Handlungsplanung	.588		.735	

N= 286; Hauptkomponentenanalyse; Varimaxrotation; KMO = .565; Bartlett Test auf Sphärizität .000;
Gesamtvarianz $\delta^2_{\text{Gesamt}} = 64\%$; Eigenwerte $\lambda = 1.64; 1.28; .918$;

Tabelle 33: Dreifaktorielle Hauptachsenanalyse (Itemsetebene)

Dreifaktorielle Struktur	h^2	Faktor 1 ($\delta^2=27,3\%$)	Faktor 2 ($\delta^2=21,3\%$)	Faktor 3 ($\delta^2=15,3\%$)
Induktion	.504	.707		
Korrelation	.434	.571		
Fakten und Meinung	.124			-.347
Deduktion	.114			.283
Zielevaluation	.303		.547	
Handlungsplanung	.305		.445	

N= 286; Hauptachsenanalyse; direktes Oblimin; KMO = .565; Bartlett Test auf Sphärizität .000;

Gesamtvarianz $\delta^2_{\text{Gesamt}} = 64\%*$; Eigenwerte $\lambda = 1.64; 1.28; .918; .843$;

* Die Angabe einer Gesamtvarianz lässt aufgrund der Berücksichtigung der Interkorrelation der Faktoren keinen aussagekräftigen Schluss zu.

Tabelle 34: Konfirmatorische Faktorenanalyse zur Prüfung der Validität der Faktormodelle

Faktormodell (N=286)	χ^2	df	p	CMIN/df	RMSEA	CFI	NFI	TLI
Itemsetebene								
Einfaktorielles Modell (Subtestebene)	27,7	9	.001	3,08	.086	.791	.734	.651
Zweifaktorielles Modell (Itemsetebene)	10,260	8	.247	1,28	.031	.975	.902	.953
Einzelitemebene								
Dreifaktorielles Modell (a)	181,04	64	.000	2,83	.080	.567	.499	.385
Dreifaktorielles Modell (b)	108,15	63	.000	1,72	.050	.833	.701	.759
Vierfaktorielles Modell (a)	93,164	59	.003	1,58	.045	.874	.742	.805
Vierfaktorielles Modell (b)	103,915	59	.000	1,76	.052	.834	.713	.744
Vierfaktorielles Modell (c)	85,92	59	.013	1,46	.040	.90	.726	.846
Fünffaktorielles Modell	69,22	55	.094	1,26	.030	.947	.880	.913
Sechsfaktorielles Modell	65,50	50	.070	1,31	.033	.943	.819	.896

Tabelle 35: Erläuterung der konfirmatorisch überprüften Faktormodelle

Modell	Faktor 1	Faktor 2	Faktor 3	Faktor 4	Faktor 5	Faktor 6
Einfaktoriell	Situationsanalyse	Zielevaluation, Handlungsplanung				
Zweifaktoriell	Deduktion, Induktion, Korrelation, Fakten u. Meinung untersch.	Handlungsplanung, Zielevaluation				
Dreifaktoriell (a)	Situationsanalyse	Handlungsplanung	Zielevaluation			
Dreifaktoriell (b)	Induktion, Korrelation	Deduktion, Fakten u. Meinung untersch.	Handlungsplanung und Zielevaluation			
Vierfaktoriell (a)	Deduktion	Fakten u. Meinung untersch.	Induktion, Korrelation	Handlungsplanung, Zielevaluation		
Vierfaktoriell (b)	Deduktion, Fakten u. Meinung untersch.	Induktion, Korrelation	Zielevaluation	Handlungsplanung		
Vierfaktoriell (c)	Deduktion, Fakten u. Meinung untersch.	Induktion	Korrelation	Handlungsplanung, Zielevaluation		
Fünffaktoriell	Deduktion	Fakten u. Meinung untersch.	Induktion	Korrelation	Handlungsplanung, Zielevaluation	
Sechsfaktoriell	Deduktion	Fakten u. Meinung untersch.	Induktion	Korrelation	Zielevaluation	Handlungsplanung

Tabelle 36: Hauptkomponentenanalyse Testintelligenz und Fächerübergreifende Problemlösekompetenzen

Dreifaktorielle Struktur	h²	Faktor 1 ($\delta^2=23,5\%$)	Faktor 2 ($\delta^2=18,9\%$)	Faktor 3 ($\delta^2=15\%$)
Induktion	.637	.786		
Korrelation	.650	.792		
Fakten und Meinung	.379			-.394
Deduktion	.574			.710
Zielevaluation	.530		.715	
Handlungsplanung	.614		.781	
Testintelligenz (APM)	.630			.665

N= 286; Hauptkomponentenanalyse; Varimaxrotation; KMO = .567; Bartlett Test auf Sphärizität .000;
Gesamtvarianz $\delta^2_{\text{Gesamt}} = 57,3\%$; Eigenwerte $\lambda = 1.64; 1.32; 1.04$;

Tabelle 37: Hauptkomponentenanalyse Testintelligenz und Fächerübergreifende Problemlösekompetenzen

Vierfaktorielle Struktur	h^2	Faktor 1 ($\delta^2=23,4\%$)	Faktor 2 ($\delta^2=18,9\%$)	Faktor 3 ($\delta^2=15\%$)	Faktor 4 ($\delta^2=13\%$)
Induktion	.738	.845			
Korrelation	.669	.792			
Fakten und Meinung	.803			.892	
Deduktion	.573			-.412	.565
Zielevaluation	.733		.842		
Handlungsplanung	.616		.734		
Testintelligenz (APM)	.781				.854

N= 286; Hauptkomponentenanalyse; Varimaxrotation; MSA = .567; Bartlett Test auf Sphärizität .000;
Gesamtvarianz $\delta^2_{\text{Gesamt}} = 70,2\%$; Eigenwerte $\lambda = 1.64; 1.33; 1.04; .843$;

Tabelle 38: Hauptachsenanalyse Testintelligenz und Fächerübergreifende Problemlösekompetenzen

Vierfaktorielle Struktur	h^2	Faktor 1 $\delta^2=23,4\%$	Faktor 2 $\delta^2=18,9\%$	Faktor 3 $\delta^2=15\%$	Faktor 4 $\delta^2=13\%$
Induktion	.437	.653			
Korrelation	.447	.648			
Fakten und Meinung	.201				.447
Deduktion	.203			.299	-.323
Zielevaluation	.244		.487		
Handlungsplanung	.345		.650		
Testintelligenz (APM)	.154			.376	

N= 286; Hauptachsenanalyse; Direktes Oblimin; MSA = .567; Bartlett Test auf Sphärizität .000;
Gesamtvarianz $\delta^2_{\text{Gesamt}} = 70,2\%$; Eigenwerte $\lambda = 1.64; 1.33; 1.04; .908$

Tabelle 39: Domänengebundenheit fächerübergreifender Problemlösekompetenzen

Variablenset (z-transformierte Werte)	M*	SD	T	p	d $\frac{\mu_1 - \mu_2}{\delta}$	Signifikanz Mann-Whitney-U-Test	Signifikanz Welch-Test
Induktives Schließen							
Geisteswissenschaften	3,25	.85	-1.945	.053	.24	.049	.054
Sozialwissenschaften (Min 0, Max 4)	3,44	.79	(df= 284)				
Korrelation und Kausalität							
Geisteswissenschaften	3,27	.73	-2,032	.043	.24	.040	.044
Sozialwissenschaften (Min 0, Max 4)	3,44	.70	(df=284)				
Deduktives Schließen							
Geisteswissenschaften	4,12	.81	1.668	.096	.19	.102	.097
Sozialwissenschaften (Min 0, Max 6)	3,96	.81	(df=284)				
Fakten und Meinung unterscheiden							
Geisteswissenschaften	2,97	.71	-1.033**	.303	.13	.302	.303
Sozialwissenschaften (Min 0, Max 4)	3,05	.61	(df=254,831)				
Zielevaluation							
Geisteswissenschaften	2,83	.97	.106	.916	0	.802	.917
Sozialwissenschaften (Min 0, Max 4)	2,83	.86	(df=284)				
Handlungsplanung							
Geisteswissenschaften	3,44	2.1	-2,243	.026	.25	.044	.025
Sozialwissenschaften (Min 0, Max 10)	4,01	2.2	(df=284)				
Geisteswissenschaften N= 130							
Sozialwissenschaften N= 156							
* Mittelwerte auf Basis nicht transformierter Werte (min. und max. Bepunktung ist jeweils angegeben).							
** konservativerer Test aufgrund Varianzheterogenität.							

Tabelle 40: Mittelwertvergleich über die Semester hinweg

Variablenset	Disziplin	1. Semester		3. Semester		6. Semester	
		M	SD	M	SD	M	SD
Induktives Schließen	Geisteswissenschaften	1,32	,85	1,19	,83	1,22	,87
	Sozialwissenschaften (Min 0, Max 4)	1,33	,81	1,67	,66	1,41	,83
Mittelwertvergleich*		F=,002 (df=1), p=.96		F=7,01 (df=1), p=.01		F=1,32 (df=1), p=.25	
Korrelation und Kausalität	Geisteswissenschaften	1,23	,72	1,20	,67	1,43	,75
	Sozialwissenschaften (Min 0, Max 4)	1,25	,74	1,51	,78	1,63	,83
Mittelwertvergleich*		F=,001 (df=1), p=.98		F=5,54 (df=1), p=.02		F=1,97 (df=1), p=.16	
Deduktives Schließen	Geisteswissenschaften	1,15	,90	1,20	,70	1,02	,77
	Sozialwissenschaften (Min 0, Max 6)	1,25	,72	,77	,78	,77	,83
Mittelwertvergleich*		F=,386 (df=1), p=.54		F=5,60 (df=1), p=.02		F=2,51 (df=1), p=.12	
Fakten und Meinung unterscheiden	Geisteswissenschaften	,96	,62	,84	,78	1,07	,71
	Sozialwissenschaften (Min 0, Max 4)	1,01	,62	1,10	,68	1,05	,55
Mittelwertvergleich*		F=,198 (df=1), p=.66		F=2,29 (df=1), p=.14		F=,009 (df=1), p=.93	
Zielevaluation	Geisteswissenschaften	2,95	,85	2,63	1,07	2,82	1,04
	Sozialwissenschaften (Min 0, Max 4)	2,89	,83	2,80	,86	2,73	,90
Mittelwertvergleich*		F=,142 (df=1), p=.71		F=,545 (df=1), p=.43		F=,192 (df=1), p=.66	
Handlungsplanung	Geisteswissenschaften	3,40	1,78	2,87	2,11	3,87	2,29
	Sozialwissenschaften (Min 0, Max 10)	3,72	2,12	4,43	2,38	4,04	2,22
Mittelwertvergleich*		F=,772 (df=1), p=.38		F=8,26 (df=1), p=.01		F=138 (df=1), p=.71	

Geisteswissenschaften N= 130
Sozialwissenschaften N= 156
*Mittelwertvergleich auf Basis der z-transformierten Werte

Fragenkomplex II

Tabelle 41: Mittelwertvergleich zwischen Semestern in Bezug auf ihre Überzeugungen

	\bar{x} (SD)			d^* $\frac{\mu_1 - \mu_2}{\delta}$	ANOVA**
	Erstsemester	Drittsemester	Sechstsemester		
Transfer	2,75 (,51)	2,71 (,56)	2,55 (,52)	d=.4	F=3,82 (df=2) p<.05
Sicherheit	3,22 (,38)	3,38 (,43)	3,44 (,38)	d=.6	F=7,91 (df=2) p<.001
Quelle	2,58 (,42)	2,90 (,56)	2,91 (,59)	d=.6	F=11,55 (df=2) p<.001

*bei einem Vergleich der Erst- und Sechstsemester

**nicht-parametrische Tests bestätigen die Signifikanzen auf gleichem Niveau (Kruskal-Wallis)

N=260

Tabelle 42: Mittelwertvergleich zwischen Disziplinen in Bezug auf ihre Überzeugungen

	\bar{x} (SD)		d $\frac{\mu_1 - \mu_2}{\delta}$	T-Test*
	Sozialwissenschaftler	Geisteswissenschaftler		
Transfer	2,78 (,48)	2,53 (,57)	d=.4	T=-3,84 (df=259) p<.001
Sicherheit	3,41 (,38)	3,26 (,41)	d=.4	T=-3,86 (df=259) p<.05
Quelle	2,83 (,55)	2,73 (,54)	-	T=-1,45 (df=260) p>.05

*nicht-parametrische Tests bestätigen die Signifikanzen auf gleichem Niveau (Mann-Whitney-U)

N=260

Tabelle 43: korrelative Zusammenhänge

	1	2	3	4	5
(1) Selbstwirksamkeit	-	r= .21**	r= .23**	r= .15**	r= .34**
(2) Studieninteresse	r= .21**	-	r= .06	r= -.07	r= .38**
(3) Sicherheit	r= .23**	r= .06	-	r= .51**	r= .03
(4) Quelle	r= .15**	r= -.07	r= .51**	-	r= -.08
(5) Transfer	r= .34**	r= .38**	r= .03	r= -.08	-

N=261; ** Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant. * Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Tabelle 44: Kennwerte Selbstwirksamkeit, Studieninteresse, epistemologische Überzeugungen nach Semester und Disziplin

	Semester 1	Semester 3	Semester 6	Geisteswissenschaften	Sozialwissenschaften
Selbstwirksamkeit	2,99 (.37)	3,02 (.44)	3,05 (.41)	2,96 (.41)	3,07 (.40)
Studieninteresse	3,21 (.47)	2,99 (.62)	3,02 (.57)	3,01 (.58)	3,15 (.53)
Sicherheit	3,23 (.38)	3,38 (.43)	3,44 (.38)	3,26 (.41)	3,41 (.38)
Quelle	2,58 (.42)	2,90 (.56)	2,91 (.59)	2,73 (.54)	2,83 (.55)
Transfer	2,75 (.51)	2,71 (.56)	2,55 (.53)	2,53 (.57)	2,78 (.48)

Aufgrund der besseren Vergleichbarkeit werden trotz Ordinalskalierung arithmetische Mittelwerte und Standardabweichung statt Median und Spannweite angegeben.

N=261

Tabelle 45: Regressionsanalysen zur Vorhersage epistemologischer Überzeugungen

Kriterium	B	Standardfehler B	Beta	T	p
Transfer					
Prädiktor Studieninteresse	.306	.054	.318	5,63	.000
Prädiktor Selbstwirksamkeit	.359	.075	.271	4,80	.000
Konstante	.635	.254		2,21	.013

F=34,457; df = 2/258; R²=.211; R²_{korr}= .205; p= .000

Tabelle 46: Regressionsanalysen zur Vorhersage epistemologischer Überzeugungen

Kriterium	B	Standardfehler B	Beta	T	p
Sicherheit					
Prädiktor Quelle	.353	.039	.481	9,00	.000
Prädiktor Selbstwirksamkeit	.162	.053	.163	3,05	.000
Konstante	1,873	.181		10,330	.000

F= 50,309; df= 2/258; R²= .281; R²_{korr}= .275; p= .003

Tabelle 47: Regressionsanalysen zur Vorhersage epistemologischer Überzeugungen

Kriterium Quelle	B	Standardfehler B	Beta	T	p
Prädiktor Sicherheit	.686	.073	.505	9,406	.000
Konstante	.486	.246		1,979	.049

$F= 88,482$; $df= 1/259$; $R^2 =.255$; $R^2_{\text{korrr}}= .252$; $p= .000$

Tabelle 48: Ordinale Regression zur Vorhersage epistemologischer Überzeugungen

Kriterium Transfer	Chi-Quadrat	Cox und Snell	Nagelkerke
Prädiktor Studieninteresse			
Prädiktor Selbstwirksamkeit	109,061 (df=33)**	.342	.345

Kriterium Sicherheit	Chi-Quadrat	Cox und Snell	Nagelkerke
Prädiktor Quelle			
Prädiktor Selbstwirksamkeit	131,474 (df=34)**	.396	.400

Kriterium Quelle	Chi-Quadrat	Cox und Snell	Nagelkerke
Prädiktor Sicherheit	87.422 (df=15)**	.284	.286

** $p < .000$

Tabelle 49: Korrelative Zusammenhänge zwischen personalen Variablen und fächerübergreifender Problemlösekompetenz

	Situations- analyse	Ziel- evaluation	Handlungs- planung	Gesamttest- ergebnis	Induktion	Korrelation	Deduktion	Fakten und Meinung
Sicherheit	$r=.16$ $p=.01$	$r=.05$ $p=.43$	$r=.16$ $p=.01$	$r=.20$ $p=.001$	$r=.18$ $p=.004$	$r=.30$ $p=.00$	$r=-.21$ $p=.00$	$r=.08$ $p=.19$
Quelle	$r=.08$ $p=.22$	$r=.05$ $p=.41$	$r=.12$ $p=.05$	$r=.14$ $p=.02$	$r=.12$ $p=.05$	$r=.17$ $p=.01$	$r=-.19$ $p=.00$	$r=.08$ $p=.20$
Transfer	$r=.01$ $p=.93$	$r=.04$ $p=.51$	$r=.01$ $p=.85$	$r=.02$ $p=.71$	$r=-.02$ $p=.73$	$r=-.01$ $p=.83$	$r=.03$ $p=.62$	$r=.02$ $p=.77$
Selbst- wirksamkeit	$r=.08$ $p=.19$	$r=-.07$ $p=.24$	$r=.11$ $p=.08$	$r=.09$ $p=.13$	$r=.04$ $p=.58$	$r=.06$ $p=.32$	$r=-.08$ $p=.19$	$r=.18$ $p=.00$
Interesse	$r=-.09$ $p=.12$	$r=.09$ $p=.12$	$r=-.08$ $p=.18$	$r=-.08$ $p=.20$	$r=.03$ $p=.54$	$r=-.04$ $p=.52$	$r=-.11$ $p=.08$	$r=-.11$ $p=.09$

Tabelle 50: Verbesserung der Vorhersageleistung

Vorhersage der Testleistung nach Kenntnis ...	Subtest Situationsanalyse	Subtest Zielevaluation	Subtest Handlungsplanung	Gesamttestergebnis
... Sicherheit	$\gamma = .133$ $p = .012$	$\gamma = .035$ $p = .566$	$\gamma = .136$ $p = .013$	$\gamma = .124$ $p = .015$
... Quelle	$\gamma = .078$ $p = .151$	$\gamma = .059$ $p = .302$	$\gamma = .123$ $p = .017$	$\gamma = .095$ $p = .043$
... Transfer	$\gamma = -.018$ $p = .745$	$\gamma = .019$ $p = .752$	$\gamma = .003$ $p = .959$	$\gamma = .013$ $p = .795$
... Selbstwirksamkeit	$\gamma = .054$ $p = .332$	$\gamma = .099$ $p = .075$	$\gamma = .088$ $p = .088$	$\gamma = .038$ $p = .441$
... Interesse	$\gamma = -.132^*$ $p = .015$	$\gamma = .024$ $p = .699$	$\gamma = -.074$ $p = .165$	$\gamma = -.083$ $p = .090$

* Es interferiert das Semester: Studierende höherer Semester sind weniger interessiert an ihrem Studium, lösen jedoch die Testitems besser.

Tabelle 51: Regressionsanalysen zur Vorhersage der fächerübergreifenden Problemlösekompetenz (Gesamttestleistung)

Kriterium	B	Standardfehler B	Beta	T	p
Gesamttestleistung					
Prädiktor Sicherheit	.890	.290	.187	-3,006	.002
Konstante	2,944	.979		-3,009	.003
F= 9,399; df= 1/260; R ² = .035; p= .002					
	B	Standardfehler B	Beta	T	p
Prädiktor Quelle	.460	.216	.131	2,131	.034
Konstante	-1,246	.612		-2,035	.043
F= 4,542; df= 1/260; R ² = .017; p= .034					

Tabelle 52: Regressionsanalysen zur Vorhersage der fächerübergreifenden Problemlösekompetenz (Situationsanalyse)

Kriterium	B	Standardfehler B	Beta	T	p
Situationsanalyse					
Prädiktor Sicherheit	.385	.153	.154	2,511	.013
Konstante	-1,302	.516		-2,520	.000
F= 6,306; df= 1/260; R ² = .024; p= .013					

Tabelle 53: Regressionsanalysen zur Vorhersage der fächerübergreifenden Problemlösekompetenz (Handlungsplanung)

Kriterium	B	Standardfehler B	Beta	T	p
Handlungsplanung					
Prädiktor Sicherheit	.387	.151	.157	2,557	.011
Konstante	-1,271	.510		-2,494	.013
F= 6,537; df= 1/260; R²= .025; p= .011					
Prädiktor Quelle	.223	.112	.122	1,989	.048
Konstante	-.597	.318		-1,880	.061
F= 3,957; df= 1/260; R²= .015; p= .048					
<hr/>					
	B	Standardfehler B	Beta	T	p
Prädiktor Selbstwirksamkeit	.265	.152	.108	1,742	.083
Konstante	-.777	.463		-1,678	.095
F= 3,034; df= 1/259; R²= .012; p= .083					

Tabelle 54: Problemlösekompetenzen in Abhängigkeit epistemologischer Überzeugungen: Vergleich reflektierter und weniger reflektierter Studierender (Mediansplit)

Weniger reflektierte Studierende: N=101, $\bar{x}_{\text{Sicherheit}}=2,92$; SD=,23

Reflektiertere Studierende: N=123, $\bar{x}_{\text{Sicherheit}}=3,69$; SD=,18

T-Test	T (df)	p	d
Induktion	-2,457 (202,693)	.015	.4
Korrelation	-5,082 (222)	.000	.7
Handlungsplanung	-1,719 (222)	.087	.2

Tabelle 55: Regressionsanalysen zur Vorhersage der fächerübergreifenden Problemlösekompetenz (Testintelligenz)

Kriterium	B	Standardfehler B	Beta	T	p
Gesamtttestleistung					
Prädiktor Sicherheit	.840	.289	.176	2,902	.004
Prädiktor Testintelligenz	.213	.085	.151	2,495	.013
Konstante	-5,061	1,297		-3,904	.000

$F= 7,786$; $df= 2/258$; $R^2= .059$; $R^2_{\text{korr}}= .050$; $p= .001$

Tabelle 56: Regressionsanalysen zur Vorhersage der fächerübergreifenden Problemlösekompetenz (Testintelligenz)

Kriterium	B	Standardfehler B	Beta	T	p
Situationsanalyse					
Prädiktor Sicherheit	.395	.153	.143	2,343	.020
Prädiktor Testintelligenz	.092	.045	.125	2,036	.043
Konstante	10,313	1,074		-3,210	.001

$F= 7,916$; $df= 2/258$; $R^2= .038$; $R^2_{\text{korr}}= .031$; $p= .007$

Tabelle 57: Problemlösekompetenzen in Abhängigkeit epistemologischer Überzeugungen: Vergleich reflektierter und weniger reflektierter Studierender (Mediansplit)

	Subtest Situationsanalyse	Subtest Zielevaluation	Subtest Handlungsplanung	Gesamttest- ergebnis
Teilstichprobe 1:				
Testintelligenz	r= .115 p=.19	r= .03 p=.74	r= .20 p=.02	r=.23 p=.007
Teilstichprobe 2:				
Testintelligenz	r= .15 p=.10	r=.05 P=.59	r= .02 p=.79	r=.137 p=.12

Teilstichprobe 1
N= 145; APM Mittelwert = 10,62; SD =1,37
1. Semester N= 59, 3. Semester N= 38, 6. Semester N= 48
Geisteswissenschaften N= 71, Sozialwissenschaften N=74

Teilstichprobe 2
N= 141; Mittelwert APM = 10,81; SD = 1,40
1. Semester N= 55, 3. Semester N= 32, 6. Semester N= 54
Geisteswissenschaften N= 59, Sozialwissenschaften N=82

Tabelle 58: Regressionsanalysen zur Vorhersage der fächerübergreifenden Problemlösekompetenz (Testintelligenz)

Kriterium	B	Standardfehler B	Beta	T	p
Gesamttestleistung					
Prädiktor Sicherheit	1,032	.373	.231	2,762	.006
Prädiktor Testintelligenz	.304	.116	.219	2,633	.009
Konstante	2,355	1,691		1,392	.000

$F= 7,726; df= 2/129; R^2= .107; R^2_{\text{korr}}= .093; p= .009$

Tabelle 59: Regressionsanalysen zur Vorhersage der fächerübergreifenden Problemlösekompetenz (Testintelligenz)

Kriterium	B	Standardfehler B	Beta	T	p
Handlungsplanung					
Prädiktor Testintelligenz	.196	.084	.199	2,318	.022
Konstante	.360	.904		.399	.691

$F= 5,375; df= 1/130; R^2= .040; R^2_{\text{korr}}= .032; p= .022$

Tabelle 60: Regressionsanalysen zur Vorhersage der fächerübergreifenden Problemlösekompetenz (Testintelligenz)

Kriterium	B	Standardfehler B	Beta	T	p
Korrelation					
Prädiktor Sicherheit	.711	.149	.386	4,777	.000
Konstante	-1,195	.492		-2,429	.017

$F= 22,817; df= 1/130; R^2= .149; R^2_{\text{korr}}= .143; p= .000$

Tabelle 61: Strukturgleichungsmodell - personbezogene Kovariaten

Strukturmodell (N=286)	χ^2	df	p	CMIN/df	RMSEA	CFI	NFI	TLI
Modellfit	12,101	8	.147*	1,513	.040	.975	.936	.936
Minimum was achieved								
Regressionsgewichte			Regressionsgeschätzer (unstandardisiert)		Signifikanzniveau			
Studieninteresse → belief Transfer			.306		.000			
Selbstwirksamkeit → belief Transfer			.359		.000			
Selbstwirksamkeit → belief Quelle			.200		.015			
Selbstwirksamkeit → belief Sicherheit			.162		.002			
belief Quelle → belief Sicherheit			.355		.000			
belief Sicherheit → Fächerübergreifende Problemlösekompetenz			1,584		.000			

*H0: „Die Populationskovarianzmatrix hat die vom Modell behauptete bzw. implizierte Form“

Tabelle 62: Zusammenfassung der Beziehungen im Strukturmodell**Totale standardisierte Effekte
Modell 1**

	Selbstwirksamkeit	Quelle	Sicherheit	Studieninteresse
Quelle	.148			
Sicherheit	.234	.482		
Transfer	.271			.318
Fächerübergreifende Problemlösekompetenz	.048*	.098*	.203	

*entspricht den indirekten standardisierten Effekten

Fragenkomplex III

Tabelle 66: Punktwertbezogene Testleistung im Semestervergleich

Subtest	Erstsemester	Drittsemester	Sechstsemester
Situationsanalyse	4,75 (s= 1,48)	4,74 (s= 1,60)	4,79 (s= 1,65)
Induktion	1,325 (s=.825)	1,457 (s=.774)	1,324 (s=.846)
Deduktion	1,202 (s=.811)	0,957 (s=.769)	0,882 (s=.812)
Korrelation	1,228 (s=.729)	1,343 (s=.699)	1,539 (s=.685)
Fakten und Meinung	0,991 (s=.645)	0,985 (s=.732)	1,059 (s=.659)
Zielevaluation	2,92 (s= 0,84)	2,73 (s= 0,96)	2,78 (s= 0,96)
Handlungsplanung	3,57 (s= 1,97)	3,74 (s= 2,38)	3,96 (s= 2,25)
Gesamtscore	11,24 (s= 2,80)	11,21 (s= 3,44)	11,53 (s= 3,29)

Erstsemester N= 114, Drittsemester N=70, Sechstsemester N= 102

$N_{\text{ges}}=286$

Tabelle 67: ANOVA Transformierte Subtestleistung

ANOVA	F (df)	p	Signifikanz	
			Levene-Test	Welch-Test
Situationsanalyse	,045 (2/168,728)	.954	.684	.956
Zielevaluation	1,234 (2/166,538)	.308	.045	.294
Handlungsplanung	,915 (2/164,276)	.420	.149	.402
Gesamtscore	.188 (2/165,267)	.809	.572	.809

Erstsemester N=114; Drittsemester N= 70; Sechstsemester N=102

Tabelle 68: ANOVA Unterschiedsanalyse Erst-, Dritt- und Sechstsemester

Variablenset	F (df)	p	Signifikanz	
			Levene-Test	Welch-Test
Induktion	.729 (df=2)	.483	.263	.458
Deduktion	4,630 (df=2)	.011	.252	.012

Korrelation	5,080 (df=2)	.007	.836	.007
Fakten und Meinung	.369 (df=2)	.692	.234	.683
Situationsanalyse	.030 (df=2)	.971	.671	.972
Zielevaluation	1,244 (df=2/166,058)	.306	.039	.291
Handlungsplanung	.862 (df=2)	.424	.144	.407
Gesamtscore	.193 (df=2)	.825	.557	.843

Erstsemester N=114; Drittsemester N= 70; Sechstsemester N=102

Tabelle 69: χ^2 -Test und Zusammenhangsmaße Erst-, Dritt- und Sechstsemester

Variablenset	χ^2 (df)	p	η
Induktion	1,855 (df=4)	.762	.072
Deduktion	13,623 (df=6)	.034	.178
Korrelation	13,194 (df=4)	.010	.186
Fakten und Meinung	4,148 (df=4)	.386	.051
Situationsanalyse	14,895 (df=16)	.532	.015
Zielevaluation	62,568 (df=46)	.052	.091
Handlungsplanung	16,187 (df=20)	.705	.078
Gesamtscore	319,748 (df=320)	.634	.037

Erstsemester N=114; Drittsemester N= 70; Sechstsemester N=102

Tabelle 70: Unterschiedsanalyse T-Test Erst- und Sechstsemester

Variablenset	T (df)	p	Signifikanz Levene-Test	Welch-Test
Induktion	.01 (df= 214)	.993	.628	-
Deduktion	2,89 (df=214)	.004	.452	-
Korrelation	-3,22 (df=214)	.001	.582	-
Fakten und Meinung	-,779 (df=214)	.437	.913	-
Situationsanalyse	-,275 (df=214)	.784	.394	-
Zielevaluation	1,17 (df= 204,64)	.241	.030	.245
Handlungsplanung	-1,37 (df=214)	.174	.309	-
Gesamtscore	-,230 (df=214)	.818	.312	-

Erstsemester N=114; Sechstsemester N=102

Tabelle 71: χ^2 -Test und Zusammenhangsmaße Erst- und Sechstsemester

Variablensets	χ^2 (df)	p	η
Induktion	.379 (df=2)	.827	.000
Deduktion	11,397 (df=3)	.010	.194
Korrelation	12,804 (df=2)	.002	.215
Fakten und Meinung	.724 (df=2)	.696	.053
Situationsanalyse	9,376 (df=8)	.312	.019
Zielevaluation	21,101 (df=46)	.222	.080
Handlungsplanung	9,382 (df=10)	.496	.093
Gesamtscore	125,420 (df=129)	.573	.016

Erstsemester N=114; Sechstsemester N=102

Tabelle 72: χ^2 -Test und Zusammenhangsmaße auf Itemebene

Erst-, Dritt und Sechstsemester / *Erst und Sechstsemester (kursiv)*

Items	χ^2 (df)	ϕ^*	p
Gesundheit (Deduktion)	5,791 (2)	.143	.055
	<i>5,169 (1)</i>	<i>.155</i>	<i>.023</i>
Berufstätigkeit (Deduktion)	.950 (2)	.058	.622
	1,029 (1)	.069	.310
Familien- unternehmen (Deduktion)	5,319 (2)	.137	.070
	<i>4,264 (1)</i>	<i>.141</i>	<i>.023</i>
Handwerker (Induktion)	.489 (2)	.042	.779
	.112 (1)	.023	.738
Migration (Induktion)	2,504 (2)	.094	.286
	.140 (1)	.026	.708
Alkohol (Korrelation)	13,055 (2)	.214	.001
	12,408 (1)	.240	.000
Neuwagen (Korrelation)	2,577 (2)	.095	.276
	2,109 (1)	.099	.146

Erstsemester N=114; Drittsemester N= 70; Sechstsemester N=101

Erstsemester N=114; Sechstsemester N=101

* Cramers V entspricht in einer 2x3 Felder Kreuztabelle ϕ

0 Zellen haben eine erwartete Häufigkeit kleiner 5.

Tabelle 73: Odds Ratio

	falsche Lösung	korrekte Lösung	Odds Ratio
Alkohol			
Erstsemester	56 falsch	58 richtig	1,03 richtig
Drittsemester	32 falsch	38 richtig	1,19 richtig
Sechstsemester	26 falsch	74 richtig	2,85 richtig
Neuwagen			
Erstsemester	32 falsch	82 richtig	2,56 richtig
Drittsemester	14 falsch	56 richtig	4,0 richtig
Sechstsemester	20 falsch	81 richtig	4,05 richtig

Tabelle 74: Odds Ratio

	falsche Lösung	korrekte Lösung	Odds Ratio
Gesundheit			
Erstsemester	87 falsch	27 richtig	3,22 falsch
Drittsemester	60 falsch	10 richtig	6,0 falsch
Sechstsemester	98 falsch	12 richtig	8,17 falsch
Familienunternehmen			
Erstsemester	39 falsch	75 richtig	1,9 richtig
Drittsemester			
Sechstsemester	49 falsch	53 richtig	1,08 richtig
Berufstätigkeit			
Erstsemester	79 falsch	35 richtig	2,26 falsch
Drittsemester	50 falsch	20 richtig	2,5 falsch
Sechstsemester	76 falsch	25 richtig	3,04 falsch

Tabelle 75: T-Tests auf Variablensebene

Variablensets	T (df)	Signifikanz Levene Test	p	Welch Test
Situationsanalyse	-1,49 (284)	,118	,138	-
Induktion	-1,95 (284)	,105	.053	-
Deduktion	1,67 (284)	,861	.096	-
Korrelation	-1,95 (284)	,900	.052	-
Fakten und Meinung	-1,03 (284/254,871)	,032	.296	.303
Handlungsplanung	-2,24 (284)	.482	.026	
Zielevaluation	,106 (284)	.212	.916	
Gesamttestscore	-1,82 (284/243,215)	.008	.064	.070

Geisteswissenschaften N=130; Sozialwissenschaften N= 156

Tabelle 76: Nicht transformierte Testleistung im Disziplinvergleich

Subtest	\bar{x} (SD)	
	Sozialwissenschaften	Geisteswissenschaften
Situationsanalyse	4,89 (1,49)	4,62 (1,63)
Induktion	1,44 (.79)	1,25 (.85)
Deduktion	,96 (.81)	1,12 (.81)
Korrelation	1,44 (.70)	1,28 (.73)
Fakten und Meinung	1,05 (.61)	,97 (.71)
Zielevaluation	2,82 (.86)	2,83 (.97)
Handlungsplanung	4,01 (2,23)	3,44 (2,07)
Gesamttestscore	11,72 (3,13)	10,88 (3,40)

Geisteswissenschaften N=130; Sozialwissenschaften N= 156

Tabelle 77: Odds Ratio und χ^2 -Tests zum Disziplineinfluss auf Itemebene

	falsche Lösung	korrekte Lösung	Odds Ratio
Alkohol (Korrelation) $\chi^2=3,23$ (df=1) $\phi=.11$ p=.07			
Sozialwissenschaften	55	101	1,84 richtig
Geisteswissenschaften	59	70	1,18 richtig
Neuwagen (Korrelation) $\chi^2=1,27$ (df=1) $\phi=.07$ p=.26			
Sozialwissenschaften	32	124	3,9 richtig
Geisteswissenschaften	34	96	2,8 richtig
Migration (Induktion) $\chi^2=2,58$ (df=1) $\phi=.10$ p=.09			
Sozialwissenschaften	36	120	3,33 richtig
Geisteswissenschaften	41	89	2,17 richtig
Handwerker (Induktion) $\chi^2=3,27$ (df=1) $\phi=.11$ p=.07			
Geisteswissenschaften	56	74	1,32 richtig
Sozialwissenschaften	51	105	2,06 richtig
Arbeitslose (Fakten und Meinung unterscheiden) $\chi^2=5,12$ (df=1) $\phi=.13$ p=.024			
Geisteswissenschaften	41	89	2,17 richtig
Sozialwissenschaften	31	125	4,03 richtig
Raucher (Fakten und Meinung unterscheiden) $\chi^2=,43$ (df=1) $\phi=-.04$ p=.51			
Geisteswissenschaften	93	37	2,5 falsch
Sozialwissenschaften	117	39	3,0 falsch

0 Zellen haben eine erwartete Häufigkeit kleiner 5.

Tabelle 78: Richtige Platzierung der Maßnahmen bei der Handlungsplanung - Häufigkeiten

N=286	Richtige Platzierung der Maßnahmen	N (gültige %)
Arbeitslosigkeit	keinen Schritt richtig platziert	131 (41,1%)
	einen Schritt richtig platziert	90 (28,2%)
	zwei Schritte richtig platziert	17 (5,3%)
	drei Schritte richtig platziert	48 (15,0%)
Getränkehersteller	keinen Schritt richtig platziert	155 (48,6%)
	einen Schritt richtig platziert	84 (26,3%)
	zwei Schritte richtig platziert	21 (6,6%)
	drei Schritte richtig platziert	26 (8,2%)

Tabelle 79: Häufigste Lösungskombinationen beim Testitem „Arbeitslosigkeit“

Maßnahme	Anzahl der Einschränkungen (N)
BDCE	1 Einschränkung übersehen (N=19)
EBD	1 Einschränkung übersehen (N=20)
BDE	2 Einschränkung übersehen (N=16)
EDB	2 Einschränkungen übersehen (N=12)
BECD	3 Einschränkungen übersehen (N=13)
Einschränkungen sind	<ul style="list-style-type: none"> • An erster Stelle müssen Bildungsinvestitionen stehen • Stadtsäckel erlaubt Investitionen nur in eine Maßnahme (Bildung oder Lohnnebenkosten) • Nicht zwei Werbemaßnahmen nacheinander durchführen • Werbemaßnahme ist erst erlaubt, wenn Stadtsäckel leer ist

Tabelle 80: Alle Lösungskombinationen beim Testitem „Arbeitslosigkeit“

Lösungskombinationen	Häufigkeit	gültige Prozent	kummulierte Prozent
ABC	1	,3	13,5
ABCD	1	,3	13,8
ABD	2	,6	14,4
ABDC	3	,9	15,4
ABDCE	6	1,9	17,2
ABEC	1	,3	17,6
ABECD	1	,3	17,9
ABED	2	,6	18,5
ACDBE	1	,3	18,8
ADB	2	,6	19,4
ADBC	1	,3	19,7
ADBE	1	,3	20,1
ADC	1	,3	20,4
ADE	1	,3	20,7
AEBD	1	,3	21,0
AEBDC	3	,9	21,9
AEDB	2	,6	22,6
AEDBC	1	,3	22,9

AEDCB	1	,3	23,2
BACD	1	,3	23,5
BACED	1	,3	23,8
BADCE	2	,6	24,5
BADE	1	,3	24,8
BADEC	1	,3	25,1
BCDA	1	,3	25,4
BDA	1	,3	25,7
BDACE	1	,3	26,0
BDAE	1	,3	26,3
BDBE	1	,3	26,6
BDCAE	2	,6	27,3
BDCE	19	6,0	33,2
BDCEA	1	,3	33,5
BDE	6	1,9	35,4
BDEAC	1	,3	35,7
BDEC	8	2,5	38,2
BDECA	1	,3	38,6
BEAC	1	,3	38,9
BEACD	2	,6	39,5
BEAD	1	,3	39,8
BEADC	2	,6	40,4
BEC	1	,3	40,8
BECD	11	3,4	44,2
BECD A	2	,6	44,8
BED	5	1,6	46,4
BEDA	1	,3	46,7
BEDAC	1	,3	47,0
BEDC	5	1,6	48,6
BEDCA	3	,9	49,5
CABDE	1	,3	49,8
CBDA	1	,3	50,2
CBDE	1	,3	50,5
CDE	2	,6	51,1
CDEB	1	,3	51,4
D	1	,3	51,7
DABE	1	,3	52,0
DABEC	2	,6	52,7
DBA	3	,9	53,6

DBAC	2	,6	54,2
DBACE	2	,6	54,9
DBAEC	1	,3	55,2
DBCE	2	,6	55,8
DBE	10	3,1	58,9
DBEA	4	1,3	60,2
DBEAC	3	,9	61,1
DBEC	26	8,2	69,3
DBECA	3	,9	70,2
DCEB	1	,3	70,5
DCEBA	1	,3	70,8
DCEC	1	,3	71,2
DE	3	,9	72,1
DEA	1	,3	72,4
DEAB	7	2,2	74,6
DEAC	1	,3	74,9
DEB	11	3,4	78,4
DEBA	3	,9	79,3
DEBAC	2	,6	79,9
DEBC	3	,9	80,9
DEBCA	1	,3	81,2
EABD	3	,9	82,1
EABDC	1	,3	82,4
EAD	3	,9	83,4
EADB	5	1,6	85,0
EB	1	,3	85,3
EBAC	1	,3	85,6
EBACD	1	,3	85,9
EBD	6	1,9	87,8
EBDAC	1	,3	88,1
EBDC	10	3,1	91,2
EBDCA	3	,9	92,2
ECD	3	,9	93,1
ED	1	,3	93,4
EDA	2	,6	94,0
EDB	12	3,8	97,8
EDBAD	1	,3	98,1
EDBC	4	1,3	99,4
EDC	1	,3	99,7

EDCA	1	,3	100,0
Gesamt	277	100,0	

Tabelle 81: Häufigste Lösungskombinationen beim Testitem „Getränkehersteller“

Maßnahme	Anzahl der Einschränkungen (N)
ABE (13)	1 Einschränkung übersehen
ACD (15)	Distraktor übersehen
ADE (33)	
AED (12)	
CADE (13)	
CDEA (13)	
CEDA (10)	
DEA (17)	
EDA (12)	
Einschränkungen sind	<ul style="list-style-type: none"> • Anhöhenzufahrt schützt Vogelarten • Anhöhenzufahrt verringert Gewerbesteuer, • Verringerte Gewerbesteuer gefährdet die KITA • Bürger stimmen Bauprojekten erst zu, wenn KITA im Bau begriffen ist

Tabelle 82: Alle Lösungskombinationen beim Testitem „Getränkehersteller“

	Häufigkeit	Gültige Prozente	Kumulierte Prozente
AB	7	2,2	17,2
ABCD	1	,3	17,6
ABD	1	,3	17,9
ABDCE	1	,3	18,2
ABDE	2	,6	18,8
ABDEC	1	,3	19,1
ABE	10	3,1	22,3
ABEC	2	,6	22,9
ABECD	1	,3	23,2
ACB	1	,3	23,5
ACD	3	,9	24,5
			273

ACDE	12	3,8	28,2
ACED	3	,9	29,2
AD	1	,3	29,5
ADBC	1	,3	29,8
ADBCE	1	,3	30,1
ADBE	2	,6	30,7
ADCBE	1	,3	31,0
ADCE	2	,6	31,7
ADE	32	10,0	41,7
ADEC	1	,3	42,0
AEBCD	1	,3	42,3
AED	11	3,4	45,8
AEDC	1	,3	46,1
BA	7	2,2	48,3
BAC	1	,3	48,6
BACDE	1	,3	48,9
BCA	1	,3	49,2
BCE	1	,3	49,5
BDAE	1	,3	49,8
BDE	2	,6	50,5
BEA	4	1,3	51,7
BECD	1	,3	52,0
CA	1	,3	52,4
CAB	3	,9	53,3
CABDBE	1	,3	53,6
CABDE	1	,3	53,9
CABE	1	,3	54,2
CABED	2	,6	54,9
CAD	1	,3	55,2
CADBE	1	,3	55,5
CADE	13	4,1	59,6
CAE	1	,3	59,9
CAED	2	,6	60,5
CBA	2	,6	61,1
CBAE	1	,3	61,4
CBDEA	3	,9	62,4
CDA	2	,6	63,0
CDAB	1	,3	63,3
CDAE	6	1,9	65,2

CDAEB	1	,3	65,5
CDEA	13	4,1	69,6
CEAD	5	1,6	71,2
CEBAD	1	,3	71,5
CEBDA	1	,3	71,8
CEDA	10	3,1	74,9
DA	2	,6	75,5
DAE	6	1,9	77,4
DBEA	1	,3	77,7
DCA	1	,3	78,1
DEA	16	5,0	83,1
DEABC	1	,3	83,4
EAC	2	,6	84,0
EACD	3	,9	85,0
EAD	24	7,5	92,5
EADCB	1	,3	92,8
EBA	1	,3	93,1
EBAC	1	,3	93,4
EBCDA	1	,3	93,7
EBDA	1	,3	94,0
ECBDA	2	,6	94,7
ECDA	2	,6	95,3
ED	1	,3	95,6
EDA	11	3,4	99,1
EDAC	1	,3	99,4
EDCAB	1	,3	99,7
EDCBA	1	,3	100,0
Gesamt	319	100,0	

Tabelle 83: Schwierigkeit der Distraktoren

N=286		N (gültige %)
Item Arbeitslosigkeit	keinen erkannt	80 (28,0%)
$P_{i\ beide} = 20,63$	beide erkannt	59 (20,6%)
$P_{i\ distal} = 17,82$	nur distal	51 (16,0%)
$P_{i\ proximal} = 33,57$	nur proximal	96 (33,6%)
$P_{i\ mind.\ einen} = 72,03$		
Item Getränkehersteller	keinen erkannt	61 (21,3%)
$P_{i\ beide} = 36,71$	beide erkannt	105 (36,7%)
$P_{i\ distal} = 12,24$	nur distal	35 (12,2%)
$P_{i\ proximal} = 29,72$	nur proximal	85 (29,7%)
$P_{i\ mind.\ einen} = 78,67$		

Tabelle 84: Korrelative Zusammenhänge zwischen Teilkompetenzen des logischen Schließens im Rahmen der Situationsanalyse

Variablenset	Induktion	Deduktion	Fakten Meinung	Korrelation	Ziel- evaluation	Handlungs- planung
Induktion	-	-.16**	.09	.41**	-.00	-.00
Deduktion	-.16**	-	-.11	-.20**	.05	-.04
Fakten Meinung	.09	-.11	-	.17**	.08	.15**
Korrelation	.41**	-.20**	.17**	-	-.01	.13**
Ziel- evaluation	-.00	.05	.08	-.01	-	.24**
Handlungs- planung	-.00	-.04	.15**	.13**	.24**	-

**Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

*Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Tabelle 85: Unterschiede in der Testleistung der Handlungsplanung und Zielevaluation in Abhängigkeit der Testleistung bei der Situationsanalyse (Mediansplit)

T-Test	\bar{x} (SD)		T (df)	p	Signifikanz Levene-Test
	N=114	N=172			
Zielevaluation	2,76 (.97)	2,86 (.87)	,947 (284)	.344	.08
Handlungsplanung	3,43 (2,14)	3,96 (2,17)	-2,053 (284)	.041	.99
Gesamtscore	9,36 (2,87)	12,65 (2,56)	-9,86 (215,33)	.000	.04
Situationsanalyse	3,18 (.95)	5,82 (.82)			

Tabelle 86: Strukturmodell 1

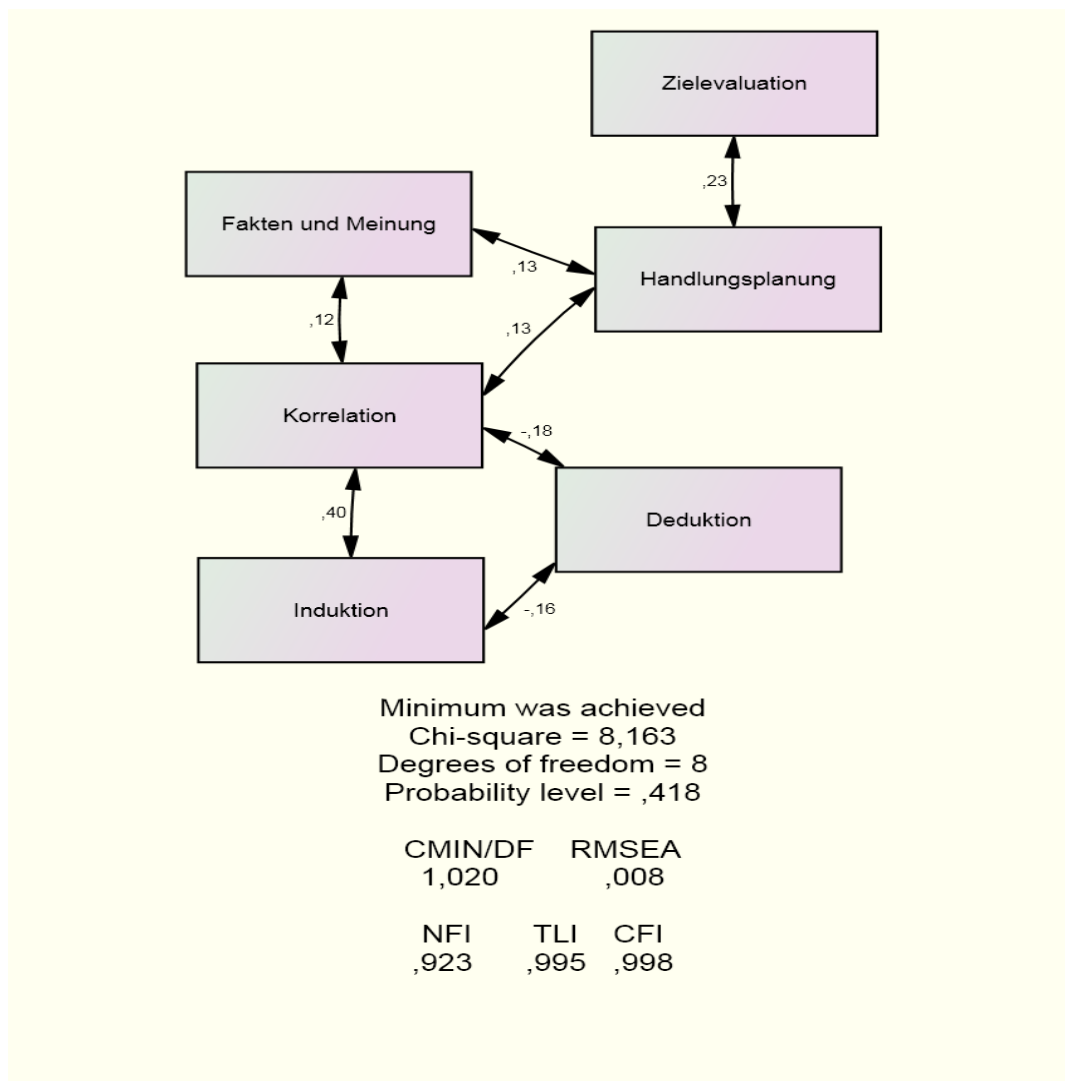


Tabelle 87: Strukturmodell 2

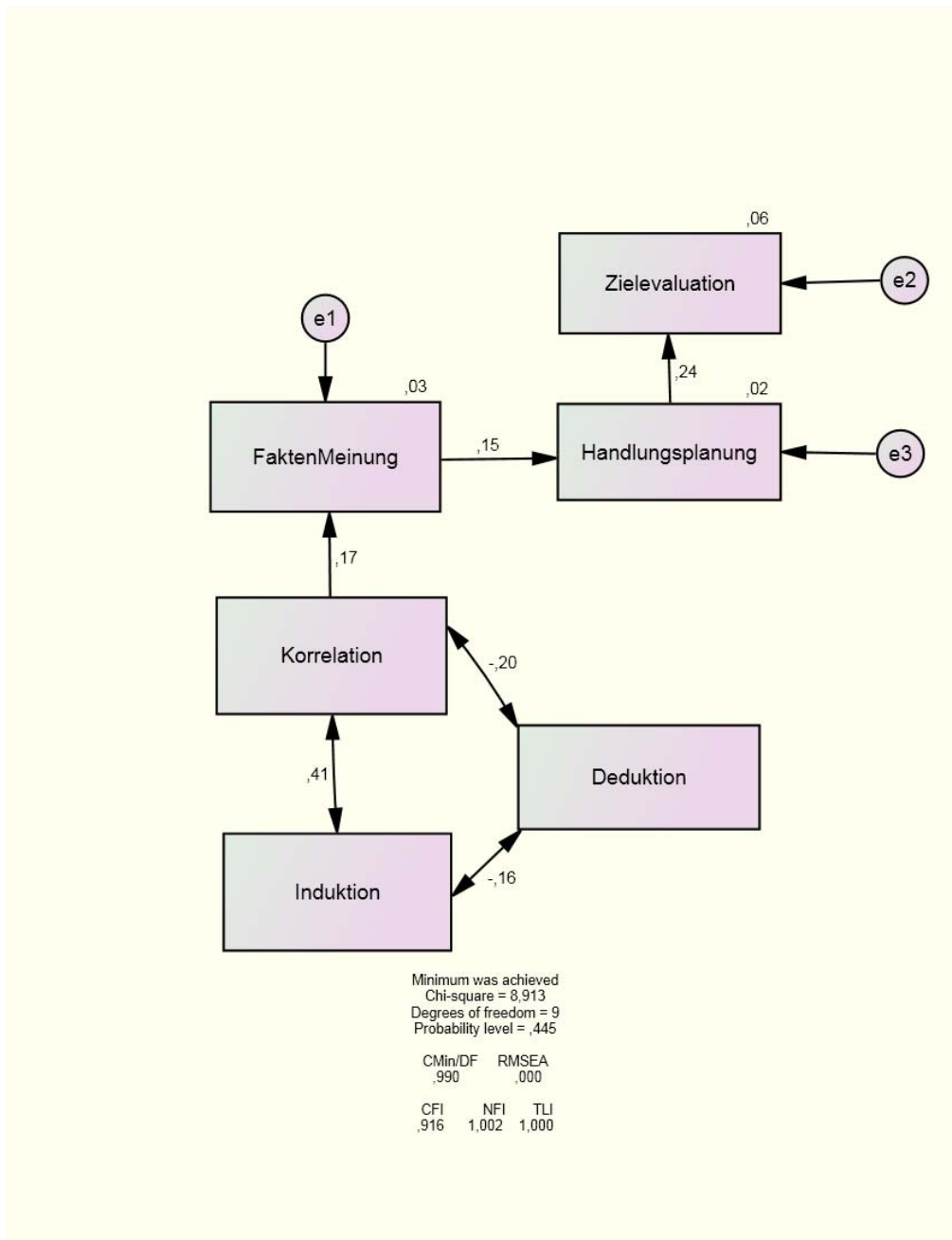
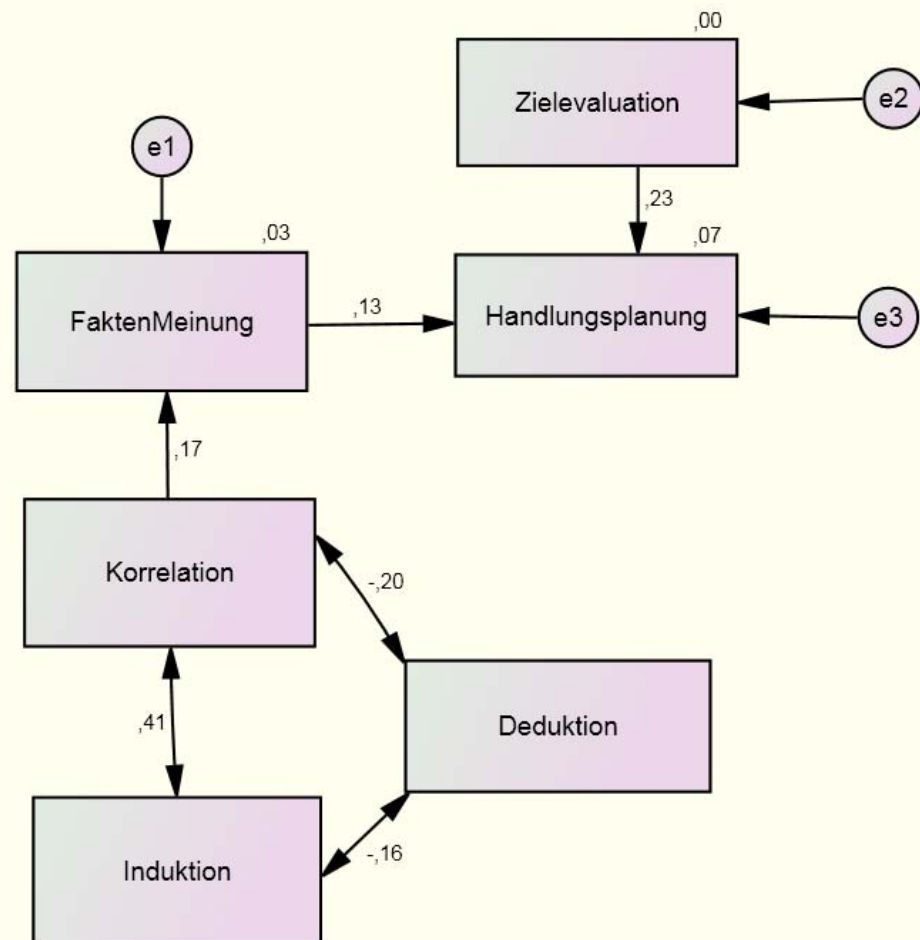


Tabelle 88: Strukturmodell 3



Minimum was achieved
 Chi-square = 10,102
 Degrees of freedom = 9
 Probability level = ,342

CMIN/DF	RMSEA
1,122	,020

CFI	TLI	NFI
,987	,970	,905

Tabelle 89: Verteilung der Semester über die Cluster

Cluster	Semester 1	Semester 3	Semester 6	N gesamt
1	34	14	19	67
2	35	19	21	75
3	26	25	44	95
4	19	12	18	49
Gesamt	114	70	102	286

Tabelle 90: Clusteranalyse

Variablenset	Cluster			
	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Induktion	-1,16	,62	,52	-,36
Deduktion	,46	,61	-,89	,17
Korrelation	-1,03	,34	,60	-,28
Fakten Meinung	,115	-,45	,55	-,55
Ziel- evaluation	,36	,57	-,02	-1,32
Handlungs- planung	,25	-,09	,35	-,89
Distanz	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1		2,36	2,77	2,42
Cluster 2	2,36		1,97	2,40
Cluster 3	2,77	1,97		2,66
Cluster 4	2,42	2,40	2,66	

Tabelle 91: Mittelwerte subjektiver Überzeugungen je nach Clusterzugehörigkeit

Cluster	Sicherheit	Quelle	Selbstwirksamkeit
1	3,22 (SD=,40)	2,70 (SD=,55)	2,99 (SD=,43)
2	3,36 (SD=,38)	2,77 (SD=,52)	2,92 (SD=,36)
3	3,50 (SD=,35)	2,93 (SD=,56)	3,13 (SD=,41)
4	3,18 (SD=,44)	2,60 (SD=,49)	2,97 (SD=,39)
Gesamt	3,35 (SD=,40)	2,78 (SD=,55)	3,02 (SD=)

Tabelle 92: Unterschiede zwischen den vier Clustern (ANOVA)

ANOVA	F (df)	p	Signifikanz	Welch-Test
			Levene-Test	
Induktion	101,96 (3/130,71)	,000	.00	.000
Deduktion	82,15 (3/133,32)	,000	.00	.000
Korrelation vs. Kausalität	59,26 (3/133,37)	,000	.00	.000
Fakten vs. Meinung	25,47 (3/140,48)	,000	.256	.000
Zielevaluation	79,64 (3/142,76)	,000	.005	.000
Handlungsplanung	43,59 (3/152,46)	,000	.003	.000
Situationsanalyse	43,65 (3)	,000	.136	.000
Gesamttest	75,44 (3)	,000	.137	.000
Skala Beliefs Sicherheit	9,33 (3)	,000	.280	.000
Skala Beliefs Quelle	4,44 (3)	,005	.626	.005
Skala Selbstwirksamkeit	4,04 (3)	,008	.475	.008
Skala Interesse	,52 (3)	,670	.286	.708
Skala Transfer	,42 (3)	,740	.267	.740

Tabelle 93: Signifikante Unterschiede zwischen den Clustern (Post-hoc-Tests, Bonferroni)

AV	Cluster	Cluster	Mittlere Differenz (I-J)	SE	p
Z-Wert Induktion	1	2	-1,79*	,11	,000
		3	-1,68*	,10	,000
		4	-,80*	,12	,000
	2	1	1,79*	,11	,000
		3	,107	,10	1,000
		4	,98*	,12	,000
	3	1	1,68*	,10	,000
		2	-,107	,10	1,000
		4	,88*	,12	,000
	4	1	,80*	,12	,000
		2	-,98*	,12	,000
		3	-,88*	,12	,000
Z-Wert Deduktion	1	2	-,14	,12	1,000
		3	1,35*	,12	,000
		4	,30	,14	,251
	2	1	,14	,12	1,000
		3	1,50*	,11	,000
		4	,44*	,14	,012
	3	1	-1,35*	,12	,000
		2	-1,50*	,11	,000
		4	-1,06*	,13	,000
	4	1	-,30	,14	,251
		2	-,44*	,14	,012
		3	1,06*	,13	,000
Z-Wert Korrelation vs. Kausalität	1	2	-1,38*	,12	,000
		3	-1,64*	,12	,000
		4	-,75*	,14	,000
	2	1	1,38*	,12	,000
		3	-,26	,11	,173
		4	,63*	,14	,000
	3	1	1,64*	,12	,000
		2	,26	,11	,173
		4	,89*	,13	,000
	4	1	,75*	,14	,000
		2	-,63*	,14	,000
		3	-,89*	,13	,000
Z-Wert Fakten vs. Meinung	1	2	,56*	,15	,001
		3	-,44*	,14	,014

		4	,66*	,16	,001
		1	-,56*	,15	,001
	2	3	-,99*	,13	,000
		4	,10	,16	1,000
		1	,43*	,14	,014
	3	2	,99*	,13	,000
		4	1,10*	,15	,000
		1	-,66*	,16	,001
	4	2	-,10	,16	1,000
		3	-1,10*	,15	,000
		2	-,21	,12	,616
	1	3	,38*	,12	,012
		4	1,67*	,14	,000
		1	,21	,12	,616
	2	3	,59*	,11	,000
		4	1,88*	,14	,000
Z-Wert Zielevaluation		1	-,38*	,12	,012
	3	2	-,59*	,11	,000
		4	1,29*	,13	,000
		1	-1,67*	,14	,000
	4	2	-1,88*	,14	,000
		3	-1,29*	,13	,000
		2	,34	,15	,154
	1	3	-,10	,14	1,000
		4	1,14*	,16	,000
		1	-,34	,15	,154
	2	3	-,44*	,13	,011
		4	,80*	,16	,000
Z-Wert Handlungsplanung		1	,10	,14	1,000
	3	2	,44*	,13	,011
		4	1,24*	,15	,000
		1	-1,14*	,16	,000
	4	2	-,80*	,16	,000
		3	-1,24*	,15	,000
		2	-1,40*	,13	,000
	1	3	-1,11*	,13	,000
		4	-,33	,15	,208
Z-Wert Situationsanalyse		1	1,40*	,13	,000
	2	3	,29	,12	,142
		4	1,07*	,15	,000
	3	1	1,11*	,13	,000

		2	-,29	,12	,142
		4	,78*	,14	,000
	4	1	,33	,15	,208
		2	-1,07*	,15	,000
		3	-,78*	,14	,000
Z-Wert Gesamttest	1	2	-1,27*	,24	,000
		3	-,83*	,23	,003
		4	2,48*	,27	,000
	2	1	1,27*	,24	,000
		3	,44	,22	,289
		4	3,76*	,26	,000
	3	1	,83*	,23	,003
		2	-,44	,22	,289
		4	3,31*	,25	,000
	4	1	-2,48*	,27	,000
		2	-3,76*	,26	,000
		3	-3,31*	,25	,000
Skala Beliefs Sicherheit	1	2	-,13	,06	,307
		3	-,27*	,06	,000
		4	,04	,07	1,000
	2	1	,13	,06	,307
		3	-,14	,06	,136
		4	,17	,07	,130
	3	1	,27*	,06	,000
		2	,14	,06	,136
		4	,31*	,07	,000
	4	1	-,04	,07	1,000
		2	-,17	,07	,130
		3	-,31*	,07	,000
Skala Beliefs Quelle	1	2	-,06	,09	1,000
		3	-,22	,08	,059
		4	,101	,10	1,000
	2	1	,06	,09	1,000
		3	-,16	,08	,334
		4	,16	,10	,695
	3	1	,22	,08	,059
		2	,16	,08	,334
		4	,33*	,10	,006
	4	1	-,10	,10	1,000
		2	-,16	,10	,695
		3	-,33*	,10	,006

Skala Selbstwirksamkeit	1	2	,07	,07	1,000
		3	-,13	,06	,274
		4	,03	,07	1,000
	2	1	-,07	,07	1,000
		3	-,20*	,06	,007
		4	-,04	,07	1,000
	3	1	,13	,06	,274
		2	,20*	,06	,007
		4	,16	,07	,163
	4	1	-,03	,07	1,000
		2	,04	,07	1,000
		3	-,16	,07	,163
Skala Interesse	1	2	-,09	,09	1,000
		3	-,06	,09	1,000
		4	-,12	,11	1,000
	2	1	,09	,09	1,000
		3	,02	,09	1,000
		4	-,03	,10	1,000
	3	1	,06	,09	1,000
		2	-,02	,09	1,000
		4	-,05	,10	1,000
	4	1	,12	,11	1,000
		2	,03	,10	1,000
		3	,05	,10	1,000
Skala Transfer	1	2	-,05	,09	1,000
		3	,03	,08	1,000
		4	,03	,10	1,000
	2	1	,05	,09	1,000
		3	,08	,08	1,000
		4	,09	,10	1,000
	3	1	-,03	,08	1,000
		2	-,08	,08	1,000
		4	,00	,10	1,000
	4	1	-,03	,10	1,000
		2	-,09	,10	1,000
		3	-,00	,10	1,000