



Philosophische Fakultät
Fachbereich Sportwissenschaft

Beurteilen sportlicher Bewegungen

Untersuchung möglicher Einflussfaktoren auf die
Beurteilungsleistung

Inauguraldissertation

zur Erlangung des akademischen Grades
eines Doktors der Philosophie (Dr. phil)

vorgelegt von

Matthias Lehner

im Wintersemester 2018/19

1. Betreuer: Prof. Dr. Gereon Berschin (Universität Passau)
2. Betreuer: Prof. Dr. Gerhard Waschler (Universität Passau)



Inhaltsverzeichnis

1	Einleitung	1
1.1	Werfen als menschliche Grundbewegung	1
1.2	Entwicklung der Forschungsfrage und Aufbau der Arbeit	3
2	Bewegungen wahrnehmen und beurteilen	7
2.1	Klärung des Untersuchungsgegenstandes	10
2.1.1	Begriffsverständnis und Abgrenzungen	10
2.1.2	Kompetenzverständnis	12
2.1.3	Bewegungsbeurteilung als diagnostische Kompetenz	17
2.1.3.1	Diagnostische Kompetenz in der Bildungsforschung	17
2.1.3.2	Beurteilungskompetenz im KMK-Modell	19
2.1.3.3	Fachdidaktische Diagnosekompetenz im Sport	22
2.1.3.3.1	Morphologische Betrachtungsweise als Mittel der Bewegungsbeobachtung	24
2.1.3.3.2	Limitierende Faktoren diagnostischer Prozesse im Sport	26
2.1.3.3.3	Einflussfaktoren auf die Bewegungsbeurteilung	27
2.1.3.3.4	Hierarchisches Ebenenmodell der Bewegungsanalyse	29
2.1.3.3.5	Methodische Arbeit mit den Meinelschen Bewegungsmerkmalen	31
2.2	Einflussfaktoren auf die Bewegungsbeobachtung	33
2.2.1	Dynamische Sehschärfe (dynamic visual acuity)	35
2.2.1.1	Physiologische Grundlagen	35
2.2.1.2	Synchronoptisches Sehvermögen	40
2.2.1.3	Messbarkeit und Trainierbarkeit	41
2.2.1.4	Schlussfolgerungen für die Messung der Beurteilungsleistung	42
2.2.2	Sportspezifische Sehstrategien	43
2.2.2.1	Visual Task Analysis der Schlagwurfbewegung	44
2.2.2.2	Wahl des Aufmerksamkeitsfokus	45
2.2.2.2.1	Eye-Tracking als Indikator für die Beobachtungsqualität?	45
2.2.2.2.2	Beobachtungsstrategien	49
2.2.2.2.3	Variablen der Aufmerksamkeitsfokussierung	53
2.2.2.2.4	Zusammenfassung	56
2.2.2.3	Gestaltung der Beobachtungssituation und Beobachterposition	57

2.2.2.4	Beobachtungswiederholung.....	58
2.3	Einflussfaktoren auf die Bewegungsbeurteilung.....	60
2.3.1	Bewegungsvorstellung	61
2.3.1.1	Begriffsbestimmung	62
2.3.1.1.1	Unterteilung der Bewegungsvorstellung	63
2.3.1.1.2	Verhältnis von kinästhetischer und visueller Bewegungsvorstellung	66
2.3.1.2	Neurophysiologische Grundlagen	67
2.3.1.3	Entstehen von Bewegungsvorstellungen	70
2.3.1.3.1	Das Information-Processing Model of Imagery nach Farah	71
2.3.1.3.2	Der sprachlich-symbolische Ansatz.....	72
2.3.1.3.3	Der räumlich-bildhafte Ansatz	73
2.3.1.3.4	Der PETTLEP-Ansatz nach Holmes und Collins	73
2.3.1.4	Operationalisierung der Bewegungsvorstellung	74
2.3.1.4.1	Selbsteinschätzungsfragebögen	75
2.3.1.4.2	Mentale Chronometrie	76
2.3.1.4.3	Mentale Rotation	79
2.3.1.4.4	Movement Imagery Specific Test.....	79
2.3.1.4.5	Zusammenhang verschiedener Testverfahren	80
2.3.2	Sollwertpräzision und Analogie-Schluss.....	83
2.3.2.1	Empfehlungen für die Formulierung von critical features	84
2.3.2.2	Technik-Cues	86
2.3.2.3	Einsatz von Analysebögen	87
2.3.2.4	Adressatenbezug	90
2.3.2.5	Zusammenfassung.....	91
2.3.3	Bewegungserfahrung (Eigenrealisation)	92
2.4	Schlussfolgerungen für die methodische Konzeption dieser Studie	95
3	Fragestellung und Hypothesenbildung.....	99

4	Methodik	104
4.1	Testpersonen	106
4.2	Fachwissen über die Ausführung von Wurfbewegungen	106
4.2.1	Wissen über den räumlich-zeitlichen Verlauf der Wurfbewegung.....	110
4.2.1.1	Sollwertbestimmung der Wurfbewegung.....	111
4.2.1.2	Testverfahren für das räumlich-zeitliche Wissen	116
4.2.2	Biomechanik.....	117
4.2.2.1	Sachanalyse des biomechanischen Wissens	117
4.2.2.2	Testverfahren für das biomechanische Wissen	123
4.2.3	Fehlerbilder	125
4.2.3.1	Sachanalyse des Fehlerwissens	125
4.2.3.2	Testverfahren für das Fehlerwissen	127
4.2.4	Visuelle Bewegungsvorstellung.....	127
4.2.5	Gütekriterien	128
4.2.5.1	Objektivität	128
4.2.5.2	Reliabilität.....	130
4.2.5.3	Validität	131
4.3	Test zur Kontrolle der Bewegungsvorstellung (TKBV)	133
4.3.1	Testdurchführung.....	135
4.3.2	Gütekriterien	136
4.3.3	Exkurs: Test zur Bewegungsbeobachtung (BeBo) als Erweiterung des TKBV	138
4.4	Eigenrealisation	139
4.4.1	Qualitative Detailanalyse	140
4.4.2	Biomechanische Datenerhebung	141
4.4.2.1	Biomechanische Analyse der Speerwurfbewegung	141
4.4.2.1.1	Bewegungstempo: Anlaufgeschwindigkeit.....	142
4.4.2.1.2	Bewegungsrhythmus: Impuls- und Stemmschritt	143
4.4.2.1.3	Bewegungsumfang: Hüft- und Schulterrotation in der Wurfauslage und Wurfarmellbogen.....	145
4.4.2.1.4	Bewegungskopplung: Beschleunigungsverlauf der Körperteilbewegungen	147
4.4.2.1.5	Bewegungspräzision: Kniewinkel des Stemmschritts in der Wurfauslage.....	150
4.4.2.2	Biomechanische Datenerhebung mittels Dartfish	151
4.4.3	Verwendung der biomechanischen Daten.....	157

4.5 Beurteilungsfähigkeit der leichtathletischen Schlagwurfbewegung	158
4.5.1 Beschreibung der Beurteilungsmerkmale	158
4.5.2 Konzeption eines Beurteilungsrasters.....	161
4.5.3 Beurteilungsaufgaben	162
4.5.4 Inter-Coder-Reliabilität	164
4.5.5 Gütekriterien	168
4.5.5.1 Objektivität	168
4.5.5.2 Reliabilität.....	169
4.5.5.3 Validität	170
4.6 Vorerfahrung im Beurteilen sportlicher Bewegungen	171
4.7 Datenauswertung	172
4.7.1 Die Item-Response-Theorie.....	172
4.7.1.1 Das dichotome Rasch-Modell.....	173
4.7.1.2 Zentrale Annahmen und Eigenschaften des Rasch-Modells	175
4.7.1.3 Prüfung der Modellauswahl	178
4.7.1.4 Statistische Prüfung der Modellgültigkeit	180
4.7.1.4.1 Nicht-parametrische Überprüfung mittels quasi-exakter Testverfahren	181
4.7.1.4.2 Parametrische Berechnungen.....	187
4.7.1.5 Differential Item Functioning (DIF).....	192
4.7.1.6 Trennschärfe.....	194
4.7.1.7 Schätzung der Personenparameter	195
4.7.1.8 Das ordinale Rasch-Modell (Partial-Credit-Modell)	197
4.7.2 Codierung	203
4.7.2.1 Auswertung nach dem dichotomen Rasch-Modell	203
4.7.2.2 Auswertung nach dem ordinalen Rasch-Modell	204
4.7.3 Normalverteilung	205
4.7.4 Korrelations- und Regressionsanalyse	208
4.7.5 Nicht-parametrischer Gruppenvergleich	210

5	Ergebnisse	212
5.1	Zusammenhänge der möglichen Einflussfaktoren mit der Beurteilungsleistung	212
5.1.1	Messung der Beurteilungsleistung	212
5.1.2	Fachwissen	216
5.1.2.1	Zusammenfassung der Testergebnisse	216
5.1.2.2	Zusammenhang zwischen Fachwissen und Beurteilungsleistung	218
5.1.3	Fähigkeit zum Ist-Sollwert-Vergleich	219
5.1.3.1	Zusammenfassung der Testergebnisse	219
5.1.3.2	Zusammenhang zwischen TKBV und Beurteilungsleistung	220
5.1.4	Beurteilungserfahrung	222
5.1.4.1	Zusammenfassung der Testergebnisse	222
5.1.4.2	Zusammenhang zwischen der Beurteilungserfahrung und der Beurteilungsleistung	223
5.1.5	Eigenrealisation	228
5.1.6	Zusammenfassung: Mögliche Wirkungen der Einflussfaktoren auf die Beurteilungsleistung	229
5.1.6.1	Einflussfaktoren auf die Fehleranalyse	231
5.1.6.2	Einflussfaktoren auf die Beurteilung einzelner Bewegungsmerkmale	232
5.1.6.3	Einflussfaktoren auf die Unterscheidung der Qualitätsstufen	233
5.1.6.4	Einflussfaktoren auf die Beurteilung des Bewegungsvergleichs	233
5.1.6.5	Einflussfaktoren auf die Beurteilung der synchronen Gesamtanalyse	234
5.1.6.6	Bildung gemeinsamer Beurteilungsindices	235
5.2	Detaillierte Auswertung und Analyse der Testverfahren	236
5.2.1	Fachwissenstest	236
5.2.1.1	Bestimmung der Personen- und Itemparameter	236
5.2.1.2	Prüfung der Modellauswahl	240
5.2.1.3	Auswertungsalternativen	244
5.2.2	Test zur Kontrolle der Bewegungsvorstellungsfähigkeit (TKBV)	246
5.2.2.1	Item- und Personenparameter des TKBV	247
5.2.2.2	Prüfung der Modellauswahl des TKBV	248
5.2.2.3	Exkurs: Ergebnisse des Bewegungsbeobachtungstests	249
5.2.2.3.1	Modellprüfung des Bewegungsbeobachtungstests	249
5.2.2.3.2	Item- und Personenparameter des Bewegungsbeobachtungstests	250
5.2.2.3.3	Zusammenhang des TKBV mit dem Bewegungsbeobachtungstest	251

5.2.3	Messverfahren der Eigenrealisation	252
5.2.4	Test zur Erfassung der Beurteilungsleistung der Schlagwurfbewegung	256
5.2.4.1	Teil A: Fähigkeit zur Unterscheidung der Qualitätsstufen.....	257
5.2.4.2	Teil B: Fähigkeit zur Einschätzung der Qualitätsstufen	260
5.2.4.3	Teil C: Fähigkeit zum Vergleich von Bewegungsmerkmalen	262
5.2.4.4	Teil D: Fähigkeit zum Erkennen fehlerhafter Bewegungsmerkmale	263
5.2.4.5	Teil E: Fähigkeit zur synchronen Gesamtanalyse der Bewegungsmerkmale	265
5.2.4.6	Modellprüfungen.....	268
6	Diskussion	271
6.1	Messung der Beurteilungsleistung	271
6.2	Bewegungsbeurteilung und Fachwissen	273
6.3	Bewegungsbeurteilung und Eigenrealisation	275
6.4	Bewegungsbeurteilung und Fähigkeit zum Ist-Sollwert-Vergleich	278
6.5	Bewegungsbeurteilung und Beurteilungserfahrung.....	279
6.6	Praxisrelevanz und Kritik der Ergebnisse.....	280
7	Ausblick.....	285
8	Appendix.....	288
9	Literaturverzeichnis	325

Abbildungsverzeichnis

<i>Abb. 1:</i> Einflussgrößen auf die beim Bewegungssehen zu erbringende Beurteilungsleistung (nach Tidow, 1993, S. 17)	7
<i>Abb. 2:</i> Einordnung der visuellen Beurteilung von Fremdbewegungen in den Kontext der allgemeinen Bewegungsbeurteilung (modifiziert nach Gralla et al., 1998, S. 166)	11
<i>Abb. 3:</i> Bewegungssehen im Sport und korrespondierende Beobachtungsaufgaben bei Einnahme der Außenperspektive (nach Tidow, 1996b, S. 248)	12
<i>Abb. 4:</i> Entwicklung von Fertigkeit, Fähigkeit und Kompetenz durch Tätigkeit nach Frey (2006, S. 32).....	15
<i>Abb. 5:</i> Diagnostische Kompetenz im Professionsmodell nach Krauss et al. (2004).....	18
<i>Abb. 6:</i> Diagnostische Fähigkeiten im COACTIV-Modell nach Brunner et al. (2011, S. 217).....	19
<i>Abb. 7:</i> Kompetenzerwerb im Rahmen der Lehrerbildung (nach Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung, 2013, S. 3)	20
<i>Abb. 8:</i> Aus den Kompetenzen abgeleitete Studieninhalte nach dem KMK-Modell (Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung, 2013, S. 48)	21
<i>Abb. 9:</i> Kennzeichnung von Bewegung nach phänomenaler und struktureller Ebene (Baumann, 1986, S. 50)	24
<i>Abb. 10:</i> Bewegungsbeurteilung beeinflussende Aspekte (nach Hartmann, 1999, S. 107).....	28
<i>Abb. 11:</i> Hierarchisches Modell allgemeiner und spezifischer Struktur- und Güteaspekte (Hartmann, 1999, S. 113)	31
<i>Abb. 12:</i> Duale Auswertung dynamischer und statischer Sehinformationen über parietale (magnozelluläre) und temporale (parvozelluläre) Verarbeitung (Ehrenstein & Jendrusch, 2008, S. 10).....	36
<i>Abb. 13:</i> Möglichkeiten der Bewegungswahrnehmung: afferentes Bewegungssehen über retinale Bildwanderung (a) und efferentes Bewegungssehen über Auge-Kopf-Bewegungen (Neumaier, 1979, S. 45).....	36
<i>Abb. 14:</i> Koordination von Augen- und Kopfbewegung bei reflektorischer horizontaler Blickbewegung auf einen plötzlich im rechten Gesichtsfeld auftauchenden Lichtreiz (am Beispiel des Rhesusaffens) (Grüsser & Grüsser-Cornhels, 1997, S. 279)	37
<i>Abb. 15:</i> Veränderung der Sehschärfe auf unterschiedlichen Retinaregionen (nach Grüsser & Grüsser-Cornhels, 1997, S. 300)	38
<i>Abb. 16:</i> Größe des Sehwinkels bei synchronoptischen Beobachtungsaufgaben (Gralla, 1996, S. 307)	41
<i>Abb. 17:</i> Verteilung der Fixationsorte bei der Beobachtung von Tennisaufschlägen mit unterschiedlicher Aufgabenstellung (a: Antizipation der Schlagrichtung, b: Identifikation von Bewegungsfehlern) (nach Neumaier (1988, S. 270))	47
<i>Abb. 18:</i> Intuitive Sehstrategie-Varianten bei Realbeobachtung und geforderter komplexer Beurteilung der Kugelstoßbewegung (nach Tidow, Koch und Marées (1984, S. 133)).....	48

<i>Abb. 19:</i> Sehstrategie-Vorgabe mit nur drei Phasenelement-Beurteilungen (Tidow, 1996b, S. 261).....	48
<i>Abb. 20:</i> Modell für die Durchführung einer Bewegungsbeobachtung (nach Barrett, 1984, S. 23).....	60
<i>Abb. 21:</i> Unterteilung der Vorstellungsmodalitäten (eigene Darstellung)	64
<i>Abb. 22:</i> Unterteilung der Bewegungsvorstellung (modifiziert nach Munzert und Reiser (2003, S. 221))	65
<i>Abb. 23:</i> Neuroanatomische Lokalisation wichtiger motorischer Zentren des menschlichen Gehirns (Medial-Schnitt, modifiziert nach Trepel (2004))	69
<i>Abb. 24:</i> Veranschaulichung der Interaktionsprozesse, die zur Bewegungsvorstellung führen (modifiziert nach Munzert, 2001b, S. 45).....	71
<i>Abb. 25:</i> Wahrnehmungs- und Vorstellungsprozess nach Farah (1984, S. 250)	71
<i>Abb. 26:</i> Beispiel einer Aufgabe aus dem TKBV mit der Bedingung Wiedererkennen (Schott, 2013, S. 667)	79
<i>Abb. 27:</i> Beispiele für Fragen und Antwortmöglichkeiten des MIST (Moreau et al., 2010, S. 99).....	80
<i>Abb. 28:</i> Bildkartenmaterial des Bildkarten-Auswahl-Tests für die Große Körperwelle (Daugis et al., 1989, S. 50): die quer bzw. diagonal verlaufende Linie verbindet die korrekten Ausführungspositionen	82
<i>Abb. 29:</i> Analysebogen-Grundmodell der O'Brien-Technik im Kugelstoßen (nach Tidow, 1996b, S. 257)	89
<i>Abb. 30:</i> Beziehung möglicher Einflussfaktoren mit der Beurteilungsfähigkeit sportlicher Bewegungen.....	101
<i>Abb. 31:</i> Konstruktionsschema für motorische Tests (Bös, 2017b, S. 812).....	105
<i>Abb. 32:</i> Modell professioneller Handlungskompetenz – Professionswissen (nach Baumert und Kunter (2006, S. 482)).....	106
<i>Abb. 33:</i> Strukturmodell für das Fachwissen Wurfbeziehung.....	108
<i>Abb. 34:</i> Fuß- und Speerweg auf den letzten vier Schritten eines 94,44 m-Wurfes von oben betrachtet (Strüder et al., 2016, S. 671).....	119
<i>Abb. 35:</i> Beispielhafter zeitlicher Verlauf der Speerbeschleunigung und der Bodenreaktionskräfte ab dem Impulsschritt (Strüder et al. (2016, S. 676) modifiziert nach Adamczewski (1995))	119
<i>Abb. 36:</i> Impulsschritt mit Verdeutlichung von Rücklage und Achsenparallelität (Strüder et al., 2016, S. 672).....	120
<i>Abb. 37:</i> Teilkörper- und Speergeschwindigkeiten am Beispiel des Speerwurfs einer Siebenkämpferin (Killing, 2015, S. 189).....	121
<i>Abb. 38:</i> Speerwerfer mit ausgeprägter Wurfseitenfreiheit aus der Rückperspektive (Strüder et al., 2016, S. 679).....	122
<i>Abb. 39:</i> Testaufgabe Visuelle Bewegungsvorstellung (Bilder modifiziert nach Haberkorn und Plaß (1992, S. 150f.))	128

<i>Abb. 40:</i> Lösungsmöglichkeiten für Aufgabe 1 des TKBV-Teil 1 gemäß dem Original-Testmanual von Schott (die Nummerierung stellt eine Modifikation zur besseren Zuordnung der Lösungsmöglichkeiten im Rahmen dieser Studie dar)	135
<i>Abb. 41:</i> Phasenstruktur der letzten beiden Schritte nach Bradshaw et al. (2012, S. 42)	143
<i>Abb. 42:</i> Messzeitpunkte t_2 , t_1 und t_0 nach Campos et al. (2004, S. 48)	144
<i>Abb. 43:</i> Messwerte für die Preparatory Phase (blau) und die Final Delivery Phase (gelb) in Sekunden (nach Campos et al. (2004, S. 49))	145
<i>Abb. 44:</i> Visualisierung des acceleration path zu den Zeitpunkten t_2 (links) und t_1 (rechts) nach Campos et al. (2004)	146
<i>Abb. 45:</i> Ellbogenwinkel beim Abwurf (Murakami et al., 2006, S. 77, Figure 10)	147
<i>Abb. 46:</i> Geschwindigkeitsverläufe beim Wurf von Parvianen (89.52) nach Campos et al. (2004, S. 50)	147
<i>Abb. 47:</i> Bestimmung der Abwurfhöhe nach Murakami et al. (2006, S. 74, Figure 5)	149
<i>Abb. 48:</i> Beugung des Oberkörpers beim Abwurf (Murakami et al., 2006, S. 79, Figure 14)	149
<i>Abb. 49:</i> Bestimmung von Oberkörperwinkel und Wurfarmwinkel nach Bradshaw et al. (2012, S. 43)	150
<i>Abb. 50:</i> Bestimmung des Stemmbeinwinkels nach Murakami et al. (2006, S. 78, Fig. 13)	151
<i>Abb. 51:</i> Dartfish-Analyseschritt 1	154
<i>Abb. 52:</i> Dartfish-Analyseschritt 2	154
<i>Abb. 53:</i> Dartfish-Analyseschritt 3	155
<i>Abb. 54:</i> Dartfish-Analyseschritt 4	156
<i>Abb. 55:</i> Dartfish-Analyseschritt 5	156
<i>Abb. 56:</i> Aufgabencharakteristische Kurve (ICC) für eine Aufgabe mit Schwierigkeit $\beta_j = 6$ (Strobl, 2012, S. 10)	174
<i>Abb. 57:</i> Aufgabencharakteristische Kurven (ICCs) für mehrere Aufgaben eines Tests (Strobl, 2012, S. 11)	175
<i>Abb. 58:</i> ICC für die Aufgaben des Fachwissenstests Teil A	177
<i>Abb. 59:</i> ICC-Kurven eines 2PL-Modells (Bartolucci et al., 2016, S. 105)	179
<i>Abb. 60:</i> Der Stichprobenraum Ω_{rc} enthält 7×13 Matrizen (schwarze Punkte). Von der Ausgangsmatrix (weißer Punkt) wird nach 8 Schritten (burn-in-Wert, schwarze Linie) die erste effektive Matrix (blauer Punkt) erreicht. Dann werden noch zwei weitere effektive Matrizen mit der Schrittweite 4 erzeugt ($n_{eff} = 3$). Alle Punkte entlang der schwarzen Linie sind generierte Matrizen, die aber nicht in die Berechnung von Test-Statistiken einfließen.	182
<i>Abb. 61:</i> Beurteilungstest Teil D5 und D6 mit Splitkriterium Gender	188
<i>Abb. 62:</i> DIF-Plot für Teil A des Fachwissenstests	189
<i>Abb. 63:</i> Vergleich der Leistungskurven von Jungen und Mädchen für zwei unterschiedliche Items (Blot 3 und Blot 35) (Bond & Fox, 2015, S. 107)	193
<i>Abb. 64:</i> Person-Item-Map für Teil C des Beurteilungstests	196

<i>Abb. 65:</i> Beispiel für ein fünfkategoriales Item mit seinen vier Schwellenfunktionen und den ermittelten Schwellenwerten $\tau_{j1} = -1.0$, $\tau_{j2} = 0.5$, $\tau_{j3} = 1.0$ und $\tau_{j4} = 2.0$ (Rost, 2004, S. 200).....	199
<i>Abb. 66:</i> Die Kategorienfunktionen des fünfkategoriellen Items von Abb. 65 (Rost, 2004, S. 200).....	199
<i>Abb. 67:</i> Ein vierkategoriales Item mit ungeordneten Schnittpunkten: Schnittpunkt $p(1)-p(2)$ liegt weiter links als Schnittpunkt $p(0)-p(1)$ (Rost, 2004, S. 198).....	200
<i>Abb. 68:</i> Idealtypische Kategorienfunktionen für ein dreikategoriales Item (Rost, 2004, S. 197).....	200
<i>Abb. 69:</i> ICC-Plot für die Items E1.2 und E1.6 des Beurteilungstests.....	201
<i>Abb. 70:</i> Person-Item-Map für Teile E1 und E2 des Beurteilungstests als ordinales Rasch-Modell.....	202
<i>Abb. 71:</i> Histogramm für die Verteilung der Personenwerte bei Teil D des Beurteilungstests.....	206
<i>Abb. 72:</i> Normalverteilungsdiagramm (links) und trendbereinigtes Q-Q-Diagramm (rechts) für Teil B des Beurteilungstests (ausgewertet nach dem ordinalen Rasch-Modell).....	206
<i>Abb. 73:</i> Boxplot für die Variable BT_B_ORM (Beurteilungstest Teil B nach dem ordinalen Rasch-Modell)	207
<i>Abb. 74:</i> Gruppiertes Boxplot für die Variable BT_B_ORM (Beurteilungstest Teil B berechnet nach dem ordinalen Rasch-Modell) nach den Stufen der Beurteilungsvorerfahrung.....	207
<i>Abb. 75:</i> Plot der Korrelation ER_ORM – B ohne Jitter (links) und mit Jitter (rechts).....	209
<i>Abb. 76:</i> Durchschnittlicher Personenwert für die Einzelaufgaben des Beurteilungstests nach dem dichotomen (DRM) und dem ordinalen Rasch-Modell (ORM).....	213
<i>Abb. 77:</i> Werteverteilung der Einzelteile des Beurteilungstests	214
<i>Abb. 78:</i> Boxplot für die Teile C, D und E des Beurteilungstests	215
<i>Abb. 79:</i> Korrelationsbeziehung zwischen den Variablen BT_D und BT_E mit Jitter (D = 3, E = 4)	216
<i>Abb. 80:</i> Histogramm für die Personenparameter der Teile A (links) und B (rechts) des Fachwissenstests mit Normalverteilungskurve.....	216
<i>Abb. 81:</i> Erreichte Punkte bei den Fragen zum Fehlerwissen über Anlauf, Abwurf und Abfangen	217
<i>Abb. 82:</i> Verteilung der Personenparameter des TKBV, wenn beide Teile als Gesamtscore ausgewertet werden	219
<i>Abb. 83:</i> Histogramm für die Personenparameter der Teile 1 (links) und 2 (rechts) des TKBV mit Normalverteilungskurve.....	220
<i>Abb. 84:</i> Zuordnung der Testpersonen zu den Stufen der Vorerfahrung in der Beurteilung sportlicher Bewegungen (N = 59).....	222
<i>Abb. 85:</i> Box-Plot des Kruskal-Wallis-Test bei unabhängigen Stichproben: Median der einzelnen Vorerfahrungsstufen für Teil A des Beurteilungstests (ausgewertet nach dem ORM)	225

<i>Abb. 86:</i> Häufigkeit der Personenparameter der Eigenrealisation mit Normalverteilungskurve.....	228
<i>Abb. 87:</i> Standardisierte Regressionskoeffizienten (β) der Prädiktoren in ihrer Wirkung auf die Beurteilungsleistung der Einzelteile des Beurteilungstests (die mit * markierten β -Werte sind auf dem 0,05-Niveau signifikant)	230
<i>Abb. 88:</i> Person-Item-Map für Teil A des Fachwissenstests	237
<i>Abb. 89:</i> Person-Item-Map für Teil B (Biomechanik)	238
<i>Abb. 90:</i> Graphischer Modelltest für Teil A des Fachwissenstests mit Splitkriterium Gender.....	242
<i>Abb. 91:</i> Person-Item-Map für Teil A und B des Fachwissenstests nach Eliminierung der leichten Items	245
<i>Abb. 92:</i> Person-Item-Map für Teil A und B des Fachwissenstests nach KPrim-Auswertung und Berechnung nach dem Partial-Credit-Modell	246
<i>Abb. 93:</i> Person-Item-Map für die 19 Items des TKBV (Teil 1 und 2).....	248
<i>Abb. 94:</i> Person-Item-Map für die 9 Items des Bewegungsbeobachtungstests	251
<i>Abb. 95:</i> Person-Item-Map für die 39 Beobachtungsdetails der Eigenrealisation (ausgewertet nach dem ordinalen Rasch-Modell)	252
<i>Abb. 96:</i> Durchschnittlich erreichte Punktzahl der einzelnen Beurteilungskriterien (er1 bis er39)	253
<i>Abb. 97:</i> Biomechanische Messwerte der Anlaufgeschwindigkeit zum Zeitpunkt t_0 (in m/s) der 59 Testpersonen	254
<i>Abb. 98:</i> Zeitliches (oben) und räumliches Verhältnis (unten) der letzten beiden Anlaufschritte	255
<i>Abb. 99:</i> Differenz der Anlaufgeschwindigkeit zwischen t_1 zu t_3	255
<i>Abb. 100:</i> Biomechanische Messwerte des Abflugwinkels in Grad mit den Intervallbereichen der Qualitätsstufen	256
<i>Abb. 101:</i> Anzahl der richtigen Einschätzungen der Qualitätsstufen für die Bewegungsmerkmale	257
<i>Abb. 102:</i> Person-Item Map nach dem DRM für Teil A des Beurteilungstests	258
<i>Abb. 103:</i> Person-Item Map nach dem ORM für Teil A des Beurteilungstests	259
<i>Abb. 104:</i> Person-Item-Map für Teil B.....	260
<i>Abb. 105:</i> Person-Item-Map für Teil B des Beurteilungstest nach dem ordinalen Rasch-Modell.....	261
<i>Abb. 106:</i> Person-Item-Map für Teil C des Beurteilungstests als dichotomes Rasch-Modell.....	262
<i>Abb. 107:</i> Person-Item-Map für Teil C (Aufgaben 1 bis 6) des Beurteilungstests als ordinales Rasch-Modell.....	263
<i>Abb. 108:</i> Person-Item Map Teil D(12) nach dem dichotomen Rasch-Modell.....	264
<i>Abb. 109:</i> Person-Item Map Teil D(12) nach dem ordinalen Rasch-Modell.....	265

<i>Abb. 110:</i> Aufgabenschwierigkeit nach Bewegungsmerkmal für Aufgaben E1 und E2 des Beurteilungstests.....	266
<i>Abb. 111:</i> Person-Item-Map für Teil E1 und E2 des Beurteilungstests als dichotomes Rasch-Modell.....	267
<i>Abb. 112:</i> Person-Item-Map für Teile E1 und E2 des Beurteilungstests als ordinale Rasch-Modell.....	268
<i>Abb. 113:</i> Task sheet für die Bewegungsbeobachtung des Überkopfwurfes nach Dunham (1994).....	288
<i>Abb. 114:</i> Anweisungen der zehn Aufgaben des TKBV (Teil 1)	305
<i>Abb. 115:</i> Anweisungen der zehn Aufgaben des TKBV (Teil 2)	306
<i>Abb. 116:</i> Testbogen für den TKBV-Teil 1 und den Bewegungsbeobachtungstest.....	309
<i>Abb. 117:</i> Auszug aus dem original TKBV-Testmanual von Schott (Seite 1)	310
<i>Abb. 118:</i> Testmanual Biomechanisch Analyse Wurfbewegung mit Dartfish.....	315
<i>Abb. 119:</i> Testbogen zur Erfassung der Beurteilungsvorerfahrung	319
<i>Abb. 120:</i> Person-Item-Map für Teil A und B des Fachwissenstests	323

Tabellenverzeichnis

Tab. 1: Übersicht zu wesentlichen Einflussgrößen auf die Bewegungsbeurteilung (nach Neumaier, 1988, S. 279)	8
Tab. 2: Kategorien des Bewegungssehens nach Neumaier und Jendrusch (1999, S. 129)	11
Tab. 3: Fragen zur Bewegungsbeobachtung der acht Beobachtungsdurchgänge nach Logan und McKinney (1977, S. 213f.)	50
Tab. 4: Beobachtungsvorlage für den Überkopfwurf mit Einteilung in drei Bewegungsphasen nach Gangstead und Beveridge (1984, S. 62).....	52
Tab. 5: Cue-basierter Beobachtungsbogen für die O'Brien-Technik im Kugelstoßen (nach Fronske, 2014, S. 368f.)	89
Tab. 6: Technikanalysebogen für die Phasen Anlauf und Speerrücknahme (Tidow, 1996a, S. 60)	112
Tab. 7: Testvignette zur Phase der Ballrückführung (schwarz gefüllte Felder sind richtig formulierte Items, unausgefüllte Felder sind falsche Antworten)	112
Tab. 8: Technikanalysebogen für die Phasen des Impulsschritts (Tidow, 1996a, S. 60).....	113
Tab. 9: Testvignette zur Phase des Impulsschrittes	113
Tab. 10: Testvignette zur Phase der Wurfauslage	114
Tab. 11: Technikanalysebogen für die Phasen des Stemmschritts und der Bogenspannung (Tidow, 1996a, S. 61)	114
Tab. 12: Technikanalysebogen für die Phasen der Schlagbewegung und des Abfangens (Tidow, 1996a).....	115
Tab. 13: Testvignetten zu den Phasen der Abwurfbewegung (4.1 und 4.2) und des Abfangens (5.1)	116
Tab. 14: Testvignetten zum biomechanischen Wissen	124
Tab. 15: Literatur über Fehlerbeschreibungen im Speerwurf und Ballwurf	125
Tab. 16: Erwartungshorizont für das Fehlerwissen der Phase Anlauf azyklisch (reduziert auf die Nennung der Fehlerbilder ohne Ursachenbeschreibungen)	126
Tab. 17: Kameraperspektiven der zehn Aufgaben des Bewegungsbeobachtungstests	139
Tab. 18: Rater des Bewegungsbeurteilungstests mit Qualifikationsnachweis	164
Tab. 19: Expertenurteile für Item A2	165
Tab. 20: Expertenurteile für Teil B des Beurteilungstests	166
Tab. 21: Expertenurteile für Teil C des Beurteilungstests	167
Tab. 22: Expertenurteile für die Aufgaben D1 und D2 des Beurteilungstests	167
Tab. 23: Expertenurteile für die Aufgaben E1 und E2 des Beurteilungstests	168
Tab. 24: Cronbach's Alpha für die Einzelteile des Beurteilungstests	169
Tab. 25: Cronbach's Alpha für die Bewegungsmerkmale des Beurteilungstests.....	170
Tab. 26: Informationstheoretische Maße für die Teile E3 und E4	180

Tab. 27: Alpha-Korrektur für die quasi-exakten Tests nach Koller et al. (2012).....	187
Tab. 28: Personenparameter des Fachwissenstests	196
Tab. 29: Experteneinschätzung der Items A3 und A4 mit Codierung der Antwortmöglichkeiten	203
Tab. 30: Expertenurteile für Teil B des Beurteilungstests	204
Tab. 31: Codierung von Einzelitems entsprechend der Übereinstimmung mit den Einschätzungen der fünf Experten (observer a, b, c, d und e)	205
Tab. 32: Korrelationskoeffizienten zwischen der Auswertung nach dem dichotomen und dem ordinalen Rasch-Modell.....	213
Tab. 33: Nichtparametrische Korrelationen zwischen den nach dem dichotomen Rasch-Modell geschätzten Personenwerten der Einzelteile des Beurteilungstests (nach Spearman)	215
Tab. 34: Korrelationen zwischen den Teilen des Beurteilungstests (ausgewertet nach dem ORM) und den Teilen des TKBV (ausgewertet nach dem DRM)	221
Tab. 35: Mittelwerte und Standardabweichung der Personenwerte der Beurteilungsaufgaben (berechnet nach dem ORM) für die einzelnen Stufen der Beurteilungsvorerfahrung	223
Tab. 36: Ergebnisse des Kruskal-Wallis-Tests für die einzelnen Teile des Beurteilungstests hinsichtlich der fünf Vorerfahrungsstufen	224
Tab. 37: Mittlerer Rang der einzelnen Vorerfahrungsstufen für die Teile des Beurteilungstests	225
Tab. 38: Asymptotische Signifikanzwerte (2-seitig) des Mann-Whitney-U-Tests für den Gruppenvergleich bezüglich verschiedener Variablen.....	226
Tab. 39: Itemschwierigkeit für die 19 Items des TKBV.....	247
Tab. 40: Itemschwierigkeit für die 9 Items des Bewegungsbeobachtungstests	250
Tab. 41: AIC- und BIC-Index für das Rasch- und das 2PL-Modell für die Einzelteile des Beurteilungstests	268
Tab. 42: Signifikanzwerte der Teststatistik T_{10} für die Splitkriterien Median, Gender und Mean für die Einzelteile des Beurteilungstests	269
Tab. 43: Ergebnisse der Teststatistiken T_{md} und T_{11}	270
Tab. 44: Fragebogen zur Prüfung des Fachwissens über die leichtathletische Schlagwurfbewegung	289
Tab. 45: Synopse der Fehlerbilder für die Phasen Anlauf (Frage 7.1) und Abwurf (Frage 7.2).....	293
Tab. 46: Erwartungshorizont für das Fehlerwissen der Phase Anlauf azyklisch	296
Tab. 47: Erwartungshorizont für das Fehlerwissen der Phase Wurfgerätrückführung	298
Tab. 48: Erwartungshorizont für das Fehlerwissen der Phase Abwurf.....	299
Tab. 49: Erwartungshorizont für das Fehlerwissen der Phase Abfangen	304
Tab. 50: Beobachtungsdetails der Wurfbewegung (B.Mw = Nummerierung des biomechanischen Messwerts aus Tab. 51)	311
Tab. 51: Biomechanische Messwerte der Dartfish-Analyse	316

Tab. 52: <i>Biomechanische Messdaten einer Testperson aus der Videoanalyse der Wurfbewegung</i>	317
Tab. 53: <i>Beurteilungsraster für die Qualitätsstufen der Bewegungsmerkmale</i>	318
Tab. 54: <i>Korrelationsmatrix zwischen der Beurteilungsleistung von Teil D und den möglichen Einflussfaktoren</i>	320
Tab. 55: <i>Korrelationsbeziehungen zwischen der Beurteilungsleistung von Teil B des Beurteilungstests und den möglichen Einflussfaktoren</i>	320
Tab. 56: <i>Korrelationsbeziehungen zwischen der Beurteilungsleistung von Teil A des Beurteilungstests und den möglichen Einflussfaktoren</i>	321
Tab. 57: <i>Korrelationsbeziehungen zwischen der Beurteilungsleistung von Teil C des Beurteilungstests und den möglichen Einflussfaktoren</i>	321
Tab. 58: <i>Korrelationsbeziehungen zwischen der Beurteilungsleistung von Teil E des Beurteilungstests und den möglichen Einflussfaktoren</i>	321
Tab. 59: <i>Likelihood-Ratio-Tabelle für Teil A und B des Fachwissenstests</i>	322
Tab. 60: <i>p-Werte der Teststatistik T4 für Teil A des Fachwissenstests</i>	322
Tab. 61: <i>p-Werte der Teststatistik T4 für Teil B des Fachwissenstests mit Splitkriterium Geschlecht</i>	323
Tab. 62: <i>Test auf Normalverteilung der Einzelteile des Beurteilungstests</i>	324

1 Einleitung

1.1 Werfen als menschliche Grundbewegung

Das Werfen stellt eine spezifisch menschliche Grundbewegung dar. Kleinkinder beginnen nach dem Erlernen des Greifens intuitiv damit, Gegenstände zu werfen. Das Schleudern von Steinen in ein Gewässer zählt bei vielen Kindern zu den ersten spannenden Tätigkeiten, in die sie sich minutenlang vertiefen können. Im Laufe der kindlichen Entwicklung werden verschiedene Wurfarten erprobt, um Geräte präzise in ein Ziel zu treffen oder möglichst weit zu werfen. In vielen unserer modernen Sportarten spielt das Werfen eine zentrale Rolle, sei es als präziser Wurf eines Baseballspielers, als kräftiger leichtathletischer Speerwurf oder als Handballwurf. Dabei hat sich der Schlagwurf als die zweckmäßigste Wurfart herausgebildet, um sowohl eine hohe Präzision als auch eine große Wurfkraft verwirklichen zu können. In einer sehr ähnlichen Form wird die Schlagbewegung auch in anderen Sportarten wie dem Schmetterschlag im Volleyball, dem Tennisaufschlag oder dem Badmintonsmash ausgeführt. Daher stellt der Schlagwurf eine wichtige sportliche Grundbewegung dar, die von Heranwachsenden in verschiedenen sportlichen Kontexten angewendet werden kann. Damit Kinder und Jugendliche das Werfen im Sportunterricht oder im Vereinstraining richtig erlernen und ihre Bewegungsausführung kontinuierlich verbessern, ist es von entscheidender Bedeutung, dass die in diesem Bereich tätigen Sportlehrkräfte oder Trainerinnen und Trainer in der Lage sind, die Bewegungsausführungen beobachten und beurteilen zu können. Daher soll im Rahmen dieser Studie untersucht werden, von welchen möglichen Faktoren die Beurteilungsleistung der Beobachter einer leichtathletischen Schlagwurfbewegung abhängt.

Auch evolutionsbiologisch stellt das Werfen eine zentrale Bewegungsform dar, welche die Entwicklung des Menschen maßgeblich beeinflusst hat. Das Werfen erleichterte dem Menschen als relativ langsamen, unbeweglichen Jäger nicht nur die Beschaffung fleischhaltiger Kost, sondern war auch für die Abwehr von natürlichen Feinden sowie das Ausführen aggressiver Handlungen gegenüber anderen Artgenossen nützlich. Im Vergleich zu den Primaten wie beispielsweise Schimpansen, welche nicht annähernd so kräftig werfen können, liegt bei den Menschen das Schultergelenk tiefer, wodurch eine Ausholbewegung anatomisch überhaupt erst möglich wird. Insofern kann man durchaus davon ausgehen, dass im Laufe der menschlichen Entwicklung durch die Notwendigkeit des Werfens anatomische Veränderungen einsetzten, welche zu immer besseren körperlichen Voraussetzungen für die Ausführung der Wurfbewegung führten. Roach und Richmond (2015, S. 107) behaupten aufgrund wissenschaftlicher Rekonstruktionsuntersuchungen fossiler Skelette, dass sich die entscheidende Wurfbio-mechanik vor ungefähr zwei Millionen Jahren mit dem Auftreten des Homo erectus entwickelte, von dem bereits angespitzte Holzspeere gefunden wurden.

Vor allem das Schultergelenk spielt dabei eine entscheidende Rolle, welches mit über 5000 Grad pro Sekunde (oder umgerechnet mehr als 750 Umdrehungen pro Minute) dabei die schnellste Bewegung ausführt, zu welcher der menschliche Körper in der Lage ist, wie Roach und Lieberman (2014, S. 2141f.) anhand von 20 untersuchten Baseballspielern herausfanden. Verantwortlich dafür ist vor allem die Speicherung elastischer Energie im Schultergelenk sowie die Übertragung kinetischer Energie von den proximalen auf die distalen Körpersegmente. Der Mensch ist also aufgrund der speziellen Anatomie des Schultergelenks, welches mit seinen Sehnen, Bändern und Muskeln wie ein Katapult wirkt, überhaupt erst in der Lage, so kraftvoll werfen zu können. Trotzdem ist der Wurf eine Ganzkörperbewegung, bei der viele Einzelbewegungen koordiniert werden müssen, um eine optimale Beschleunigung

des Wurfgeräts erreichen zu können. Dabei wird die meiste Energie durch die Vorspannung der Hüfte sowie die Bogenspannung im Oberkörper erzeugt, welche dafür verwendet wird, die elastischen Schulterelemente vorzuspannen, während Ellbogen und Handgelenk nur noch einen kleinen Beitrag zur Beschleunigung des Wurfgeräts liefern (Roach, Venkadesan, Rainbow & Lieberman, 2013, S. 483f.).

Wenn diese miteinander gekoppelten Teilbewegungen eingeschränkt werden, so sinkt nachweislich auch die Wurfbeschleunigung. Dies konnten Roach et al. (2013, S. 2142f.) in einer Untersuchung messen, indem bei Werfern die Beweglichkeit des Oberkörpers, des Schlüsselbeins, der Schulter und des Handgelenks mithilfe von Manschetten nacheinander eingeschränkt wurde und dadurch die Abwurfgeschwindigkeit deutlich sank. Es zeichnet demnach gute Werfer aus, die anatomisch zur Verfügung stehenden Bewegungsspielräume der genannten Körperteile optimal auszunutzen und so miteinander zu koppeln, dass über die gesamte Wurfbewegung hinweg das Wurfgerät maximal beschleunigt werden kann. Wenn Werferinnen und Werfer aus welchem Grund auch immer diese Teilbewegungen schlecht miteinander koordinieren und durch eine falsche Technik Beschleunigungspotentiale ungenutzt lassen, so wirkt sich dies unmittelbar auf die erreichte Wurfweite aus.

Dabei sind auch geschlechtsspezifische Unterschiede beim Werfen nicht von der Hand zu weisen. Ältere Untersuchungen aus den USA zeigten, dass bereits im Alter von vier Jahren die Wurftechnik bei Mädchen im Vergleich zu anderen motorischen Grundbewegungen (wie dem Ballfangen, dem Weitspringen oder dem Balancieren) signifikant schlechter ausgeprägt ist als bei Jungen (Thomas & French, 1985, S. 276). Hyde (2005) bestätigt, dass unabhängig vom Alter speziell der Schlagwurf die größten geschlechtsspezifischen Unterschiede in der motorischen Leistungsfähigkeit aufweist. Ebenfalls am Beispiel fünfjähriger amerikanischer Kinder wurde untersucht, dass Jungen durchschnittlich um 43 Prozent weiter werfen als Mädchen und im Bezug auf die Technikmerkmale Oberkörperrotation und Schrittgestaltung eine besser entwickelte Wurftechnik zeigten (Nelson, Thomas, Nelson & Abraham, 1986, S. 280). Der Geschlechtsunterschied wird mit zunehmendem Alter noch größer: Mit 13 Jahren beschleunigen die Jungen einen Tennisball beim Werfen auf durchschnittlich 85 Kilometer pro Stunde, während Mädchen lediglich 61 Kilometer pro Stunde erreichen.

Die Gründe dafür konnten noch nicht eindeutig nachgewiesen werden. Wenn man biologische Faktoren wie Körpergröße oder Armmuskulatur in Beziehung zur Wurfweite setzt, so beträgt der Wurfweitenunterschied zwischen Jungen und Mädchen beispielsweise nur noch 31 Prozent (Nelson et al., 1986, S. 280). Ebenso können umweltbezogene Gründe eine Rolle spielen. Der Sportpsychologe Heinz Krombholz vermutet beispielsweise, dass der Geschlechtsunterschied beim Werfen in Deutschland aufgrund der kulturellen Tatsache noch größer sein könnte, dass Kinder und Jugendliche in den USA flächendeckend mehr wurfbezogene Sportarten wie Baseball oder Football ausüben (Römer, 2012). Möglicherweise können auch evolutionsbiologische Gründe angeführt werden, weil im Laufe der menschlichen Entwicklung die Männer für das Jagen zuständig waren und somit täglich ihre Wurfähigkeit verbesserten, so dass dies zu einer dauerhaften Überlegenheit des männlichen Nervensystems beim Ausüben dieser motorischen Aktion führte. Thomas, Alderson, Thomas, Campbell und Elliott (2010, S. 440) haben hierzu auch herausgefunden, dass die Mädchen der australischen Aborigines im Bezug auf Abwurfgeschwindigkeit und Wurfweite den gleichaltrigen Jungen deutlich näher kommen als in Deutschland, Japan, Thailand oder den USA und immerhin 78 Prozent der Wurfgeschwindigkeit der Jungen erreichen. Als mögliche Ursache führen die Wissenschaftler an, dass in dieser australischen Kultur die Frauen gleichberechtigt für das Jagen zuständig waren und somit kein koordinativer Entwicklungsnachteil vorlag.

Der entscheidende Grund dafür, ob jemand gut oder schlecht werfen kann, wird aber immer noch in der Tatsache zu suchen sein, wie häufig Heranwachsende das Werfen üben und wie professionell sie in ihrem motorischen Entwicklungsverlauf betreut werden. An dieser Stelle nehmen Sportlehrkräfte oder Übungsleiter eine wichtige Rolle ein, indem sie die Bewegungsausführungen ihrer Sportlerinnen und Sportler beobachten, beurteilen und hilfreiche Rückmeldungen geben.

1.2 Entwicklung der Forschungsfrage und Aufbau der Arbeit

Es sollte demnach einem geschulten Beobachter auffallen, welche Körperteile nur unzureichend an der Wurfbeschleunigung beteiligt sind, um eine entsprechende Beurteilung des Wurfvorgangs vornehmen und dem Werfer gezielt Rückmeldung geben zu können. Wenn man bedenkt, dass zur ohnehin schon sehr schnell ablaufenden Wurfbewegung auch noch der Anlauf und die Abwurfvorbereitung hinzukommen, so wird die Schwierigkeit und Komplexität dieser Aufgabe, die zum täglichen Trainings- und Unterrichtsalltag gehört, schnell klar. Trotzdem gibt es kaum Untersuchungen darüber, wie man die Beurteilungsfähigkeit sportlicher Bewegungen messen oder verbessern kann. Schon Anfang der 1960er Jahre hat Meinel (1960, S. 138) darauf hingewiesen, dass der Bereich „Bewegungsanalyse – Bewegungsbeurteilung – Bewegungsanweisung [...] das Kernstück der motorisch-methodischen Denk- arbeit des Sportpädagogen“ darstellt. Im universitären Lehramtsstudium wird das Beurteilen von Bewegungen jedoch allenfalls beiläufig thematisiert und auch in der Referendariatsausbildung wird von den angehenden Sportlehrkräften das Benoten sportlicher Bewegungsabläufe verlangt, ohne die in diesem Bereich relativ unerfahrenen Lehramtsanwärter auf diese komplexe Aufgabe ausreichend vorzubereiten. Dies mag auch deshalb der Fall sein, weil wissenschaftlich fundierte Konzepte fehlen, von welchen Einflussfaktoren die Beurteilungsleistung sportlicher Bewegungen überhaupt abhängt. Wenn bekannt wäre, welches Wissen oder welche Fähigkeiten eine professionelle Bewegungsbeurteilung begünstigen, dann könnte diese diagnostische Kompetenz auch in der Lehramtsausbildung strukturierter vermittelt werden.

Umso wichtiger erscheint es, die in Training und Sportunterricht tätigen Lehrpersonen auf diese besondere Beurteilungssituation vorzubereiten. Daher soll im Rahmen dieser Studie am Beispiel der Schlagwurfbewegung untersucht werden, was einen guten Beurteiler auszeichnet. Ist es ein gut ausgeprägtes Fachwissen, welches dem Beobachter dazu dient, die entscheidenden Beobachtungskriterien überhaupt zu kennen, um sie dann detailliert beurteilen zu können? Oder hilft eine gute Eigenrealisation der Wurfbewegung dabei, die beobachtete Bewegung besser nachvollziehen und Abweichungen vom Idealbild zuverlässiger identifizieren zu können? Oder gibt es sogar eine bewegungsunabhängige Beobachtungsgabe, die dabei hilft, das Beobachtete gedanklich verarbeiten und mit früheren Erinnerungen vergleichen zu können? Möglicherweise stellt auch die Erfahrung, die ein Beobachter schon beim Beurteilen des Schlagwurfs oder schlagwurfähnlicher Bewegungen gesammelt hat, einen entscheidenden Einflussfaktor auf die Beurteilungsleistung dar. Da es in diesem Bereich noch keine wissenschaftlichen Untersuchungen gibt, wird ein eigenes Messinstrument, das sich an den häufig wiederkehrenden Alltagssituationen im Beurteilen sportlicher Bewegungen von Sportlehrkräften und Trainerinnen und Trainern orientiert, entwickelt. Es gilt demnach in dieser Untersuchung zu überlegen, wie die Beurteilungsleistung der Schlagwurfbewegung überhaupt messbar gemacht werden kann und

nach welchen Beurteilungskriterien sie erfolgen soll. Dafür können die Meinel'schen Bewegungsmerkmale eine Strukturierungshilfe bieten: mit Hilfe der Kategorien Bewegungsharmonie, Bewegungstempo, Bewegungsrhythmus, Bewegungsumfang, Bewegungskopplung und Bewegungspräzision soll in einem Beurteilungstest die Bewegungsqualität von Schlagwürfen anhand von Videobeispielen eingeschätzt werden. Dabei müssen verschiedene Beurteilungsaufgaben vom relativ einfachen Einschätzen einzelner Bewegungsmerkmale bis hin zur komplexen Fehleranalyse bewältigt werden (siehe Kapitel 4.5).

Um mögliche Einflussfaktoren auf die im Beurteilungstest gemessene Beurteilungsleistung zu untersuchen, kommen weitere Testverfahren zum Einsatz. Das Fachwissen der Beurteiler wird über einen selbst entwickelten Test gemessen, der sowohl Multiple-Choice-Fragen als auch Freitextfragen sowie einen Bildkartenauswahltest enthält (siehe Kapitel 4.2). Die Qualität der Eigenrealisation wird auf der Grundlage von Videoaufzeichnungen mit Hilfe biomechanischer Messwerte und qualitativ einzuschätzender Beurteilungskriterien bestimmt (siehe Kapitel 4.4). Um Aussagen über die allgemeine Bewegungsvorstellungsfähigkeit treffen zu können, kommt der Test zur Kontrolle der Bewegungsvorstellungsfähigkeit (TKBV) zum Einsatz, welcher mit videobasierten Bewegungsbeobachtungsaufgaben erweitert wird (siehe Kapitel 4.3). Die Beurteilungsvorerfahrung wird in ihren unterschiedlichen Ausprägungsstufen mithilfe eines Single-Choice-Fragebogens ermittelt (siehe Kapitel 4.6). Die zu untersuchenden Hypothesen werden aus dem Zusammenhang dieser vier Variablen mit der Beurteilungsleistung abgeleitet (siehe Kapitel 3).

Um das Thema in die aktuelle Bildungsdebatte theoretisch einzubinden, wird zu Beginn der Arbeit begründet, weshalb das Beurteilen sportlicher Bewegungen als Teil der Lehrkompetenz von Sportlehrkräften zu verstehen ist (siehe Kapitel 2.1). Ausgehend von der Abgrenzung mit den Begriffen Bewegungssehen, Bewegungswahrnehmung und Bewegungsbeobachtung wird die Bewegungsbeurteilung inhaltlich auf den Untersuchungsgegenstand dieser Studie verengt: die visuelle Beurteilung von Fremdbewegungen zur Überprüfung der Übereinstimmung mit einer Bewegungsvorschrift oder Idealbewegung. Tidow (1996b, S. 248f.) nennt die Qualitätsbewertung und die Soll-Istwert-Bestimmung als konkrete Beobachtungsaufgaben des Bewegungssehens, die in dieser Studie im späteren Beurteilungstest in verschiedenen Schwierigkeitsgraden untersucht werden. Das daraus resultierende fachdidaktische Anliegen, Fertigkeiten zu optimieren und Normabweichungen zu minimieren, deutet sich hierbei schon an, wird aber in dieser Studie nicht mehr behandelt (siehe Kapitel 2.1.1). Nach der Erörterung des dieser Studie zugrunde liegenden Kompetenzverständnisses (siehe Kapitel 2.1.2) und der Begriffsklärung der diagnostischen Kompetenz in der Bildungsforschung (siehe Kapitel 2.1.3.1) sowie im Kompetenzmodell der Kultusministerkonferenz (siehe Kapitel 2.1.3.2) wird in Kapitel 2.1.3.3 mit dem neu begründeten Begriff der fachdiaktischen Diagnosekompetenz auf die Besonderheiten des Beurteilens sportlicher Bewegungen eingegangen. Der Fokus liegt dabei auf der morphologischen Bewegungsbeobachtung, bei der die Bewegungsbeurteilung ohne technische Hilfsmittel (und somit ohne Möglichkeiten der Wiederholbarkeit oder Verlangsamung) auf der Grundlage der Informationsaufnahme mit dem visuellen und auditiven Analysator erfolgt. Die Schwierigkeit im Trainings- und Unterrichtsalltag besteht also darin, dass sehr schnell ablaufende Bewegungen während oder nach einer nur einmalig präsentierten Beobachtung beurteilt werden müssen. Zugleich sind sportliche Bewegungen wie der leichtathletische Schlagwurf so komplex, dass in einem Beobachtungsdurchgang realistischer Weise nur ein kleiner Teil der Beurteilungskriterien, die einen guten Schlagwurf auszeichnen, beobachtet werden können. Somit bieten sich für die morphologische Bewegungsanalyse eher übergeordnete Kategorien wie die oben bereits genannten Meinel'schen Bewegungsmerkmale an, die sich auch aus dem hierarchisch geordnete Ebenenmodell nach Hartmann (1999, S. 112f.) ableiten lassen (siehe Kapitel

2.1.3.3.4) und mit denen eine alle wichtigen Phasen umfassende Bewegungsbeurteilung machbar erscheint. Überlegungen zur methodischen Arbeit mit den Meinel'schen Bewegungsmerkmalen (siehe Kapitel 2.1.3.3.5) fassen die Voraussetzungen zusammen, die bei der Konstruktion eines Beurteilungstests sportlicher Bewegungen im Rahmen dieser Studie beachtet werden sollen.

Vor der Konstruktion der Testverfahren werden mögliche Einflussgrößen auf die Bewegungsbeurteilung untersucht, welche nach Tidow (1993) in je drei beobachtungs- und beurteilungsbezogene Faktoren eingeteilt werden können. In Kapitel 2.2 wird erörtert, warum es beim morphologischen Beurteilungsvorgang zu einer sinnesphysiologischen Überforderung des visuellen Systems kommt und welche Konsequenzen dies für die Messung der Beurteilungsleistung hat. Dabei wird die Bedeutung der drei beobachtungsbezogenen Faktoren der Bewegungs-Sehschärfe und Blickmotorik (siehe Kapitel 2.2.1) sowie der Sehstrategie (siehe Kapitel 2.2.2) auf die Beobachtungskompetenz umfassend diskutiert. Ehe die drei beurteilungsbezogenen Faktoren der Sollwertpräzision und der Fähigkeit zum Analogie-Schluss (siehe Kapitel 2.3.2) sowie der Bewegungserfahrung (siehe Kapitel 2.3.3) behandelt werden, wird der Prozess der Bewegungsvorstellung (siehe Kapitel 2.3.1) in seiner Bedeutung für die Bewegungsbeurteilung dargestellt, woraus wiederum wichtige Entscheidungen bei der Konstruktion der Testverfahren dieser Studie abgeleitet werden können. Nach dieser theoretischen Beschäftigung mit dem Konstrukt der Bewegungsbeurteilung lässt sich festhalten, dass die dynamische Sehschärfe und die Blickmotorik zwar ein leistungsbestimmender Faktor sein können, die Messbarkeit jedoch äußerst aufwendig und die Bedeutung für sportliche Beurteilungssituationen mitunter noch nicht endgültig geklärt ist, weshalb beide in dieser Studie unberücksichtigt bleiben müssen. Trotzdem fließen viele Erkenntnisse dieser menschlichen Wahrnehmungsproblematik in die methodischen Überlegungen zur Messung der Beurteilungsleistung ein. Die wissenschaftlichen Erkenntnisse über den Einfluss der Sehstrategie auf die Beurteilungsleistung werden ebenfalls nicht als Variable untersucht, sondern dienen dazu, die Beobachtung der Bewegungsmerkmale im Beurteilungstest möglichst effektiv und einheitlich zu gestalten. Zu diesem Zweck werden den Probandinnen und Probanden vor der Ausführung des Beurteilungstests die Beobachtungskriterien jedes Bewegungsmerkmals in einem Lehrvideo (siehe Beilage DVD, Beurteilungstest, Kapitel Bewegungsmerkmale) erklärt, um zu gewährleisten, dass die in einem Beurteilungsraster vorgegebenen Qualitätsstufen (siehe Appendix Tab. 53) nach der Bewegungsbeobachtung auch möglichst kohärent eingeschätzt werden.

In Kapitel 4 wird das methodische Vorgehen dieser Studie begründet. Die theoretischen Überlegungen zur Sollwertpräzision sowie zum Analogie-Schluss fließen dabei in das Messverfahren des Fachwissens (siehe Kapitel 4.2) mit ein, während der Einfluss der Bewegungserfahrung auf die Beurteilungsleistung in der Variable Eigenrealisation (siehe Kapitel 4.4) untersucht wird. Die Erkenntnisse über die Operationalisierung der Bewegungsvorstellung finden sich in der Modellierung der Variable Fähigkeit zum Ist-Soll-Wert-Vergleich wieder, welche mit dem Test zur Kontrolle der Bewegungsvorstellung (siehe Kapitel 4.3) gemessen werden soll. Die Datenauswertung der verwendeten Tests erfolgt hauptsächlich mittels probabilistischer Item-Response-Theorie nach dem Rasch-Modell (siehe Kapitel 4.7.1), wobei zur Modellprüfung bevorzugt die nicht-parametrische Überprüfung mittels quasi-exakter Testverfahren (siehe Kapitel 4.7.1.4.1) angewendet wird. Über die probabilistische Schätzung der Personenparameter (siehe Kapitel 4.7.1.7) können die Fähigkeiten einer Testpersonen bei den einzelnen Testverfahren bestimmt und miteinander verglichen werden, um Zusammenhänge zwischen der Beurteilungsleistung und möglichen Einflussfaktoren mittels Korrelation- und Regressionssanalysen (siehe Kapitel 4.7.4) feststellen zu können. Mit Hilfe eines nicht-parametrischen Gruppenvergleichs (siehe Kapitel

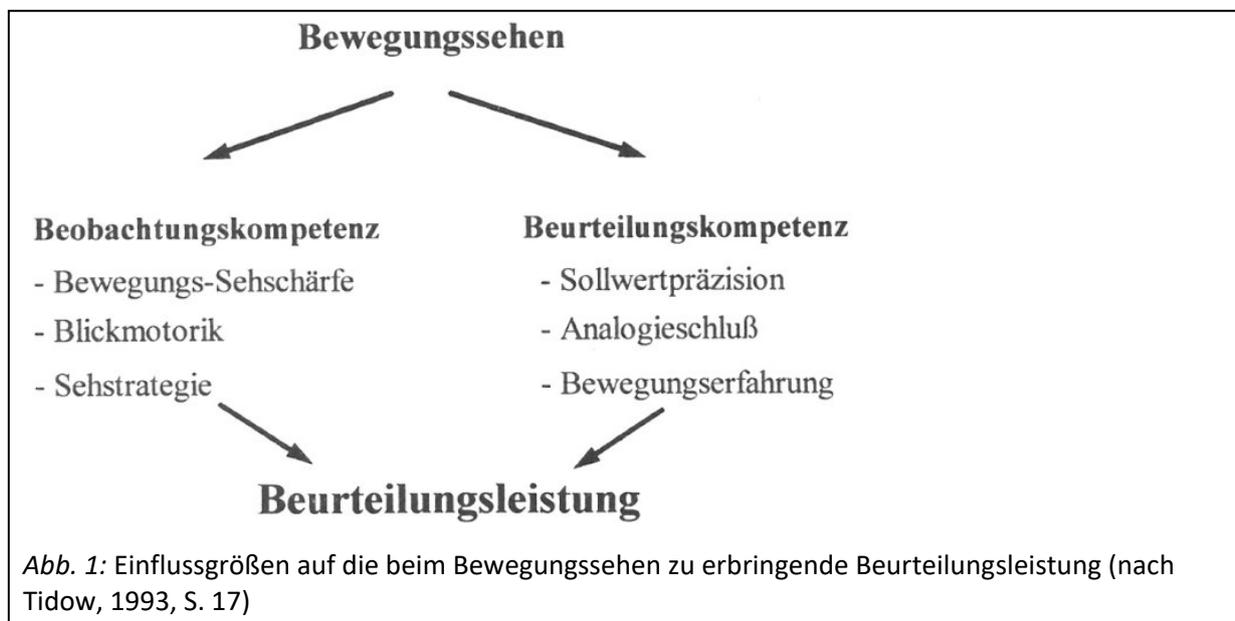
4.7.5) kann schließlich untersucht werden, ob die Beurteilungsleistung von der Spezifität der Beurteilungsvorerfahrung (siehe Kapitel 4.6) der Probandinnen und Probanden als letzten untersuchten Einflussfaktor abhängt.

In Kapitel 5 werden die Ergebnisse der Untersuchung zusammenfassend dargestellt. Die wichtigsten Erkenntnisse über die Messung der einzelnen Beurteilungsaufgaben, die Testergebnisse der einzelnen Messverfahren sowie die Zusammenhänge der möglichen Einflussfaktoren mit der Beurteilungsleistung werden gleich vorweg in Kapitel 5.1 genannt, auf dessen Grundlage dann auch die Diskussion der Ergebnisse in Kapitel 6 erfolgt. Die detaillierte Auswertung und Analyse der einzelnen Testverfahren kann dann in Kapitel 5.2 nachgelesen werden. Zur Ermittlung der Eigenrealisationsqualität der Testpersonen wurden beispielsweise 39 Beobachtungsdetails mittels Videoanalyse qualitativ und quantitativ eingeschätzt, ehe aus den Daten nach Anwendung des ordinalen Rasch-Modells ein Eigenrealisationsindex berechnet wird (siehe Kapitel 5.2.3). Auf der Grundlage der entsprechenden Fragebogenergebnisse erfolgt in Kapitel 5.1.4 eine Einordnung der Probandinnen und Probanden in fünf Vorerfahrungsstufen, welche von keinerlei Beurteilungserfahrung sportlicher Bewegungen (Stufe 1) bis hin zur Beurteilungserfahrung leichtathletischer Grundbewegungen (Stufe 5) reicht. Bei den Variablen Fachwissen (siehe Kapitel 5.2.1), allgemeine Bewegungsvorstellungsfähigkeit (siehe Kapitel 5.2.2) und Beurteilungsfähigkeit der Schlagwurfbewegung (siehe Kapitel 5.2.4) wurde zuerst die Konformität mit dem Rasch-Modell statistisch geprüft, ehe die berechneten Item- und Personenparameter kritisch betrachtet und dargestellt wurden. Da die Testverfahren zum Teil aus voneinander unabhängigen Vignetten bestehen (z. B. die fünf Teilaufgaben des Beurteilungstests), werden die Ergebnisse auch detailliert für jede einzelne Vignette erläutert. So werden bei der Auswertung des Beurteilungstests fünf Personenparameter berechnet (siehe Kapitel 5.2.4), welche den einzelnen Beurteilungsteilen entsprechen und später in Beziehung zueinander gesetzt werden können (siehe Kapitel 5.1.1) sowie detailliertere Zusammenhänge mit den möglichen Einflussfaktoren (siehe Kapitel 5.1) überprüft werden können.

In Kapitel 6 werden die Ergebnisse näher diskutiert und interpretiert, bevor in Kapitel 7 ein Ausblick gegeben wird, wie die Erkenntnisse dieser Studie in neue Fragestellungen überführt werden können.

2 Bewegungen wahrnehmen und beurteilen

In diesem Kapitel werden die theoretischen Grundlagen der Bewegungswahrnehmung und Bewegungsbeurteilung behandelt, um daraus die Fragestellung dieser Studie zu entwickeln. Wer eine Bewegung beurteilen will, muss diese zuerst möglichst umfassend im Hinblick auf die vorgegebenen Beurteilungskriterien wahrnehmen. Will man Bewegungssehen als visuell-dynamischen Beurteilungsleistung eines Außenbeobachters quantifizieren, so kann man die zur Informationsaufnahme und -speicherung benötigte **Beobachtungskompetenz** von der **Beurteilungskompetenz**, welche die gespeicherten Sehinhalte identifiziert, filtert und verarbeitet, unterscheiden. Beide Kompetenzen bedingen sich dabei gegenseitig, denn „ohne präzise und bewegungsrelevante optische Informationen nützt die beste Beurteilungskompetenz wenig, ohne ein internes Bezugssystem wiederum können korrekt angenommene Bewegungsdetails keiner objektiven und zuverlässigen Bewertung zugeführt werden“ (Tidow, 1996b, S. 242). Die zu untersuchende Beurteilungsleistung hängt demnach von der Beobachtungskompetenz und der Beurteilungskompetenz ab. Die in Abb. 1 genannten Einflussfaktoren auf die Bewegungsbeobachtung werden in den Kapiteln 2.2.1 (Bewegungs-Sehschärfe und Blickmotorik) und 2.2.2 (Sehstrategie) näher behandelt, um sinnvoll abwägen zu können, welche Erkenntnisse über die visuelle Informationsaufnahme für die Konzeption der Forschungsfrage sinnvoll sein können. Bevor die Beurteilungskompetenz als zweiter Faktor der Beurteilungsleistung über das Begriffspaar Sollwertpräzision und Analogieschluss (siehe Kapitel 2.3.2) sowie über die Bewegungserfahrung (siehe Kapitel 2.3.3) konkretisiert wird, muss der Begriff der Bewegungsvorstellung (siehe Kapitel 2.3.1) geklärt werden, denn die Fähigkeit, sich Bewegungen exakt vorstellen zu können, stellt eine Grundvoraussetzung des Beurteilens sportlicher Bewegungen dar. Aus den Erkenntnissen über die Beurteilungskompetenz können wiederum mögliche Einflussfaktoren auf die Beurteilungsleistung eruiert und Schlussfolgerungen über deren Messbarkeit gezogen werden.



Neumaier (1988, S. 278-282) nimmt ebenfalls eine Zweiteilung der Beurteilungsleistung vor, indem er die sinnesphysiologisch bedingten von den psychologisch bedingten Einflussgrößen unterscheidet, wobei erstere inhaltlich der eben genannten Beobachtungskompetenz und zweitere der Beurteilungskompetenz entsprechen. Die Konkretisierung der beiden Einflussgrößen ist dahingehend interessant,

dass die beurteilerabhängigen Faktoren helfen können, ein umfassenderes Verständnis für das Anforderungsprofil einer guten Bewegungsbeurteilung zu bekommen, während die sehobjekt- bzw. aufgabenbezogenen Faktoren Hinweise liefern können, die bei der Konstruktion eines Testverfahrens für die Beurteilungsleistung beachtet werden sollten (siehe Tab. 1). Beurteilerabhängige Begriffe wie Blickbewegungen, Fixationsorte oder dynamische Sehschärfe spielen auch im Kapitel 2.2 über die Bewegungsbeobachtung eine entscheidende Rolle. Sehobjekt- bzw. aufgabenbezogene Überlegungen über die Beschaffenheit der Bewegung (z. B. der Bewegungsstruktur oder der Bewegungsgeschwindigkeit) sowie die Darbietungsform und die gewählte Beobachterperspektive geben Anhaltspunkte, um die Beurteilungsaufgaben sinnesphysiologisch gut aufnehmen und bewerten zu können. Vor allem die psychologisch bedingten sehobjektbezogenen Faktoren wie Aufgabenstellung, Beurteilungsumfang, Informationsdichte oder Beurteilungsrichtlinien bilden wichtige Voraussetzungen für die spätere Konstruktion eines Testverfahrens, mit dem die Beurteilungsfähigkeit sportlicher Bewegungen gemessen werden soll, und werden im Kapitel 2.3 über die Bewegungsbeurteilung thematisiert, spielen aber auch bei der Behandlung sportspezifischer Sehstrategien (siehe Kapitel 2.2.2) eine wichtige Rolle.

Tab. 1: Übersicht zu wesentlichen Einflussgrößen auf die Bewegungsbeurteilung (nach Neumaier, 1988, S. 279)

Beurteilungsleistung	
1. Sinnesphysiologisch bedingte Einflussgrößen	
Beurteilerabhängig	Sehobjekt- bzw. aufgabenbezogen
1) Blickbewegungen	1) Winkelgeschwindigkeit der Bewegung
2) Fixationsorte	2) Darbietungszeit (Bewegungsstruktur)
3) Statische und dynamische Sehschärfe	3) Darbietungsform (Realsituation/Film)
4) Peripheres Sehen	4) Beobachterperspektive
5) Tiefenwahrnehmung	5) Sehwinkel
	6) Beleuchtung
2. Psychologisch bedingte Einflussgrößen	
Beurteilerabhängig	Sehobjekt- bzw. aufgabenbezogen
1) Vorinformation / Erwartung	1) Aufgabenstellung/Beurteilungsgegenstand
2) Bewegungserfahrung	2) Beurteilungsumfang
3) Beurteilungserfahrung	3) Informationsdichte
4) Kapazität des Kurzzeitspeichers	4) Beurteilungsrichtlinien
5) Psychologisch bedingte Beurteilungsfehler	

Diesen beiden Theoriekonzepten über die Beurteilungsleistung strukturell folgend beschäftigt sich Kapitel 2.2 zuerst mit dem Prozess des Bewegungssehens als absichtsvolle Beobachtung von Bewegungsmerkmalen, um die sinnesphysiologischen Voraussetzungen und Limitationen zu klären, mit denen der optische Analysator überhaupt Bewegungen wahrnehmen kann (siehe Kapitel 2.2.1.1), speziell wenn es um synchronoptische Beobachtungsvorgänge geht, wie sie in sportlichen Beurteilungssituationen häufig vorkommen (siehe Kapitel 2.2.1.2). Auch die Messbarkeit und Trainierbarkeit der dynamischen Sehschärfe (siehe Kapitel 2.2.1.3) liefert einige Anhaltspunkte, um den Beobachtungsvorgang in seiner Bedeutung für die Bewegungsbeurteilung besser verstehen zu können. Gerade weil schnell ablaufende sportliche Bewegungen wie der leichtathletische Schlagwurf so schwer zu beobachten sind, kommt den sportspezifischen Sehstrategien eine entscheidende Rolle zu. Je besser ein Beobachter auf die Beobachtungssituation vorbereitet ist, desto erfolgsversprechender wird die Informationsaufnahme während der Bewegungsbeobachtung sein. Über eine Visual Task Analysis (siehe Kapitel 2.2.2.1) kann eine erste Auseinandersetzung mit der konkret zu beobachtenden Bewegung sein, um dann den Aufmerksamkeitsfokus entsprechend der Beurteilungsaufgabe passend zu wählen (siehe Kapitel 2.2.2.2) und die Beobachtungssituation entsprechend zu gestalten (siehe Kapitel 2.2.2.3). In Kapitel 2.2.2.4 wird abschließend diskutiert, wieviele Beobachtungswiederholungen sinnvoll und notwendig sind, um eine sportliche Bewegung umfassend zu beobachten, woraus wiederum Rückschlüsse auf die spätere Konstruktion der Beurteilungskriterien gezogen werden können, mit denen die Testpersonen die Qualität von Schlagwurfbewegungen einschätzen sollen.

Zur Identifizierung möglicher Einflussfaktoren auf die Bewegungsbeurteilung ist die Beschäftigung mit dem Begriff der Bewegungsvorstellung unumgänglich, um zu klären, auf welche Weise eine Bewegung überhaupt gedanklich repräsentiert sein kann (siehe Kapitel 2.3.1.1). Nach einer Beschäftigung mit den neurophysiologischen Grundlagen (siehe Kapitel 2.3.1.2) wird das Entstehen von Bewegungsvorstellungen beleuchtet (siehe Kapitel 2.3.1.3), um davon ausgehend Möglichkeiten der Operationalisierung der Bewegungsvorstellungsfähigkeit zu diskutieren (siehe Kapitel 2.3.1.4). Diese Überlegungen sollen dazu dienen herauszufinden, ob es allgemeine kognitive Fähigkeiten gibt, welche die Ausbildung spezifischer Bewegungsvorstellung unterstützen und somit auch als mögliche Einflussfaktoren auf die Beurteilungsleistung sportlicher Bewegungen in Frage kommen. Desweiteren wird eine gelungene Bewegungsbeurteilung auch von dem Vermögen des Beurteilers abhängen, wie präzise er den Sollwert der Bewegung beschreiben und das Wissen über Zusammenhänge von Teilbewegungen für den Beurteilungsprozess nutzen kann. Um die Sollwertpräzision als möglichen Einflussfaktor valide überprüfen zu können, erfolgt zunächst eine theoretische Auseinandersetzung mit der Formulierung von sog. critical features (siehe Kapitel 2.3.2.1) und Technik-Cues (siehe Kapitel 2.3.2.2), ehe der Einsatz von Analysebögen diskutiert (siehe Kapitel 2.3.2.3) und auf die Bedeutsamkeit des Adressatenbezugs bei der Bewegungsbeurteilung hingewiesen wird (siehe Kapitel 2.3.2.4). Mit der Bewegungserfahrung des Beurteilenden (Eigenrealisation) wird der wohl augenscheinlichste Einflussfaktor auf die Bewegungsbeurteilung in Kapitel 2.3.3 dargestellt und in seiner Bedeutung für die sportliche Beurteilungsprozesse thematisiert, ehe abschließend Schlussfolgerungen für die methodische Konzeption dieser Studie gezogen werden (siehe Kapitel 2.4).

Bevor inhaltlich auf die beiden Begriffe der Bewegungsbeobachtung und Bewegungsbeurteilung näher eingegangen wird, soll zunächst der Untersuchungsgegenstand dieser Studie genauer geklärt werden. Nach der Definition und Abgrenzung der wichtigsten verwendeten Begrifflichkeiten (siehe Kapitel 2.1.1) wird das dieser Studie zugrunde liegende Kompetenzverständnis umrissen (siehe Kapitel 2.1.2). In Kapitel 2.1.3 wird dann die Bewegungsbeurteilung in den Kontext der diagnostischen Kompetenz eingeordnet, indem zunächst Überlegungen der Bildungsforschung (siehe Kapitel 2.1.3.1) dargestellt

werden und Anknüpfungspunkte zum Kompetenzverständnis der Kultusministerkonferenz hergestellt werden (siehe Kapitel 2.1.3.2). Mit dem Begriff der fachdiagnostischen Diagnosekompetenz im Sport (siehe Kapitel 2.1.3.3) wird der Versuch unternommen, die speziellen Anforderungen ernst zu nehmen, die bei diagnostischen Prozessen im Kontext des Erlernens sportlicher Bewegungen permanent vorkommen, und daraus inhaltliche und methodische Konsequenzen bezüglich der Messung der Beurteilungsleistung sportlicher Bewegungen zu ziehen.

2.1 Klärung des Untersuchungsgegenstandes

Bevor die engere Auseinandersetzung mit den Einflussgrößen auf die Beurteilungsleistung erfolgt, werden in diesem Kapitel zunächst die wichtigsten Fachbegriffe definiert und der Untersuchungsgegenstand abgegrenzt (siehe Kapitel 2.1.1). Da das Beurteilen häufig als Kernkompetenz einer Lehrkraft verstanden wird, erfolgt eine allgemeine Beschäftigung mit dem Kompetenzbegriff (siehe Kapitel 2.1.2), woraus schon grundlegende Erkenntnisse gewonnen werden können, die bei der Messung der Beurteilungsleistung beachtet werden sollen.

2.1.1 Begriffsverständnis und Abgrenzungen

Zu Beginn dieser Studie werden die Begriffe Bewegungssehen, Bewegungswahrnehmung, Bewegungsbeobachtung und Bewegungsbeurteilung definiert und voneinander abgegrenzt, um ein einheitliches Begriffsverständnis zu schaffen, ehe der Untersuchungsgegenstand genauer umschrieben wird. Der Begriff **Bewegungssehen** bezeichnet grundsätzlich „nicht nur eine Tätigkeit, sondern zugleich auch die Fähigkeit des visuellen Systems, dynamische Prozesse differenziert wahrnehmen und analysieren zu können“ (Tidow, 1993, S. 15). Neumaier (1988, S. 29) verwendet als Synonym für das Bewegungssehen den Begriff **Bewegungswahrnehmung** und versteht unter beiden den „Prozeß [sic!] der Entstehung eines (visuellen) Eindrucks einer Veränderung der Raumposition von Objekten relativ zum Wahrnehmenden“.¹ Jendrusch (2006, S. 61) bemüht sich um eine genauere Unterscheidung, indem er in Übereinstimmung mit Neumaier und Jendrusch (1999, S. 128) vom Bewegungssehen ohne bestimmte Absicht den Begriff der **Bewegungsbeobachtung** als gezieltes Bewegungssehen abgrenzt. Ist dieser Prozess darauf gerichtet, die beobachtete Bewegung (z. B. zum Zwecke der Bewegungskorrektur durch eine Lehrkraft) zu beurteilen, wird von einer visuell gestützten **Bewegungsbeurteilung** gesprochen.

Da im Rahmen dieser Arbeit kein derart differenziertes Begriffsverständnis notwendig ist, werden die Begriffe Bewegungssehen und Bewegungswahrnehmung synonym zur absichtsvollen Bewegungsbeobachtung verwendet. Die Bewegungsbeurteilung grenzt sich dahingehend von diesen Begriffen ab, dass hier speziell die Ausprägung einzelner Ausführungsmerkmale oder auch des gesamten Bewegungsablaufs in verbalen (qualitativ) oder numerischen Kategorien (quantitativ) festgestellt wird. Der Untersuchungsgegenstand der Bewegungsbeurteilung kann nach Neumaier und Jendrusch (1999, S. 129) gemäß folgender vier Kategorien (siehe Tab. 2) weiter konkretisiert werden: 1. als Wahrnehmungsperspektive wird die Außenbetrachtung einer sportlichen Bewegung eingenommen, 2. der Wahrnehmungsgegenstand entspricht einer Fremdbewegung, 3. die Wahrnehmungssituation ist auf den motorischen Lernprozess gerichtet und 4. die Funktion der visuellen Wahrnehmung liegt in der Bewegungskorrektur.

¹ Sage (1984, S. 10) definiert Wahrnehmung (perception) als „organizing and giving meaning to stimuli“.

Tab. 2: Kategorien des Bewegungssehens nach Neumaier und Jendrusch (1999, S. 129)

Kategorie	Optionen			
1. Wahrnehmungsperspektive	Aktivensicht	Außenbetrachter		
2. Wahrnehmungsgegenstand	Eigenbewegung	Fremdbewegung	Kombination	
3. Wahrnehmungssituation	Motorischer Lernprozess	Techniktraining	Wettkampf	
4. Funktion der visuellen Wahrnehmung	Orientierung im Raum	Antizipation der Ballflugbahn	Kontrolle der Eigenbewegung	Bewegungskorrektur

Will man die Funktion der visuellen Wahrnehmung noch detaillierter beschreiben, so geht es mit Gralla, Pfennig, Jendrusch, Tidow und Heck (1998, S. 166) darum, die Übereinstimmungen mit einer Idealbewegung oder Bewegungsvorschrift zu erkennen und entsprechende Bewegungsfertigkeiten zu optimieren bzw. Normabweichungen zu minimieren (Bewegungskorrektur). Die anderen in Abb. 2 genannten Aspekte spielen in dieser Studie keine Rolle.

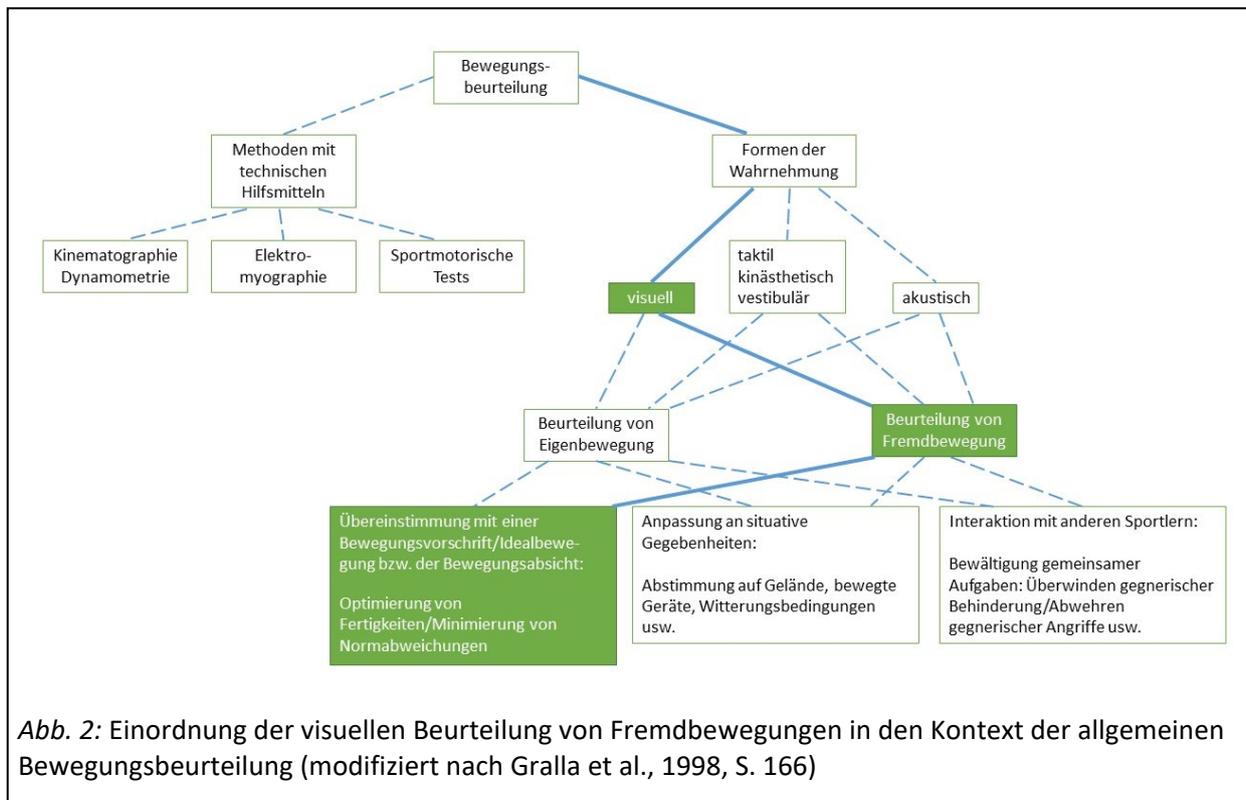
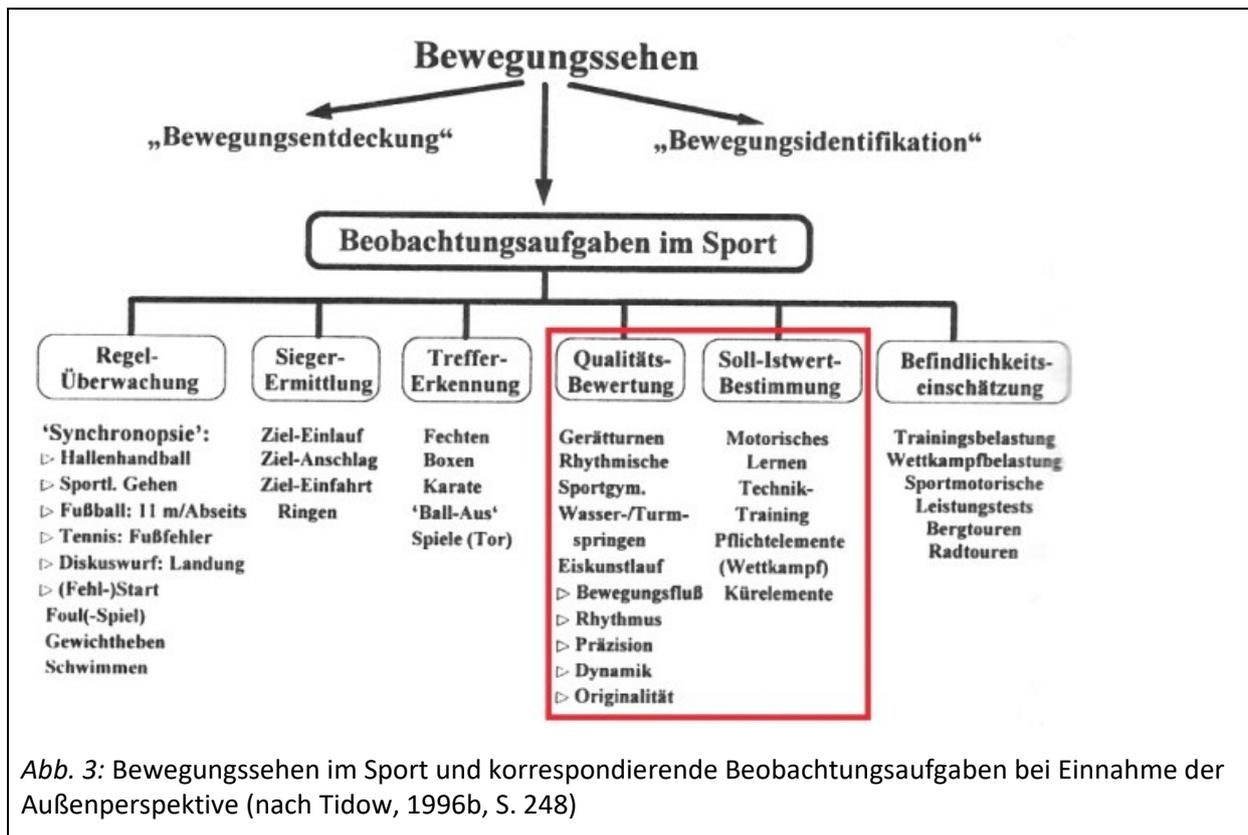


Abb. 2: Einordnung der visuellen Beurteilung von Fremdbewegungen in den Kontext der allgemeinen Bewegungsbeurteilung (modifiziert nach Gralla et al., 1998, S. 166)

Somit lässt sich der Untersuchungsgegenstand dieser Studie dahingehend formulieren, dass das visuelle Beurteilen einer Fremdbewegung aus der Außenperspektive im Mittelpunkt steht. Als zu untersuchende Bewegung wird der leichtathletische Schlagwurf gewählt, weil er eine grundlegende motorische Bewegungsform darstellt, die von Schülerinnen und Schülern im Sportunterricht gelernt und verbessert werden soll. Außerdem kommt die Schlagbewegung in anderen Sportarten wie dem Schmetterschlag im Volleyball, dem Aufschlag im Tennis oder dem Schlagwurf im Handball vor, so dass eine Übertragung der Methodik dieser Studie auf diese ähnlichen Bewegungen denkbar ist. Und zuletzt zeichnet sich die Schlagwurfbewegung durch eine gut nachvollziehbare Bewegungsstruktur aus, die in entsprechende Beobachtungsmerkmale übertragen werden kann. Aus der von Tidow (1996b, S. 248) genannten Beobachtungsaufgabe der Qualitätsbewertung werden später unter anderem die Kriterien für die Beurteilung des Schlagwurfs sowie dessen Soll-Istwert-Bestimmung abgeleitet (siehe Abb. 3, roter Rahmen). Es geht generell darum zu beurteilen, wie gut der Schlagwurf anhand vorher festgelegter Kriterien ausgeführt wurde. Dafür können die genannten Meinel'schen Bewegungsmerkmale (z. B. Umfang, Rhythmus oder Präzision) eine gute Orientierung darstellen, weil sie allgemein genug sind, um in einem Beobachtungsdurchgang wahrgenommen und beurteilt werden zu können.



2.1.2 Kompetenzverständnis

Da die Einflussgrößen auf die Beurteilungsleistung bereits mit den Begriffen Beobachtungs- und Beurteilungskompetenz bezeichnet wurden und in Kapitel 2.1.3 die Fähigkeit zur Beurteilung sportlicher Bewegungen in der bildungstheoretischen Kompetenzdebatte verortet wird, soll in diesem Kapitel das dieser Studie zugrundeliegende Kompetenzverständnis geklärt werden. Spricht man im alltäglichen Sprachgebrauch von Kompetenz, so ist damit häufig die Fähigkeit gemeint, gewisse wiederkehrende

Situationen erfolgreich zu bewältigen.² Wer hingegen inkompetent handelt, ist den Anforderungen oder der Verantwortung in einem bestimmten Bereich nicht gewachsen. Der Duden weist dem Begriff folglich die Bedeutungen Sachverstand und Fähigkeiten zu (Scholze-Stubenrecht, 2013).³ Auch wissenschaftlich betrachtet kann der Begriff Kompetenz schon auf eine lange, wenngleich je nach Fachdisziplin und Forschungsbereich sehr uneinheitliche Verwendung zurückblicken. Balz (2011) weist als Sportpädagoge beispielsweise darauf hin, dass der Kompetenzbegriff in der Bildungsforschung sowohl funktional (als angestrebtes Lernergebnis) als auch normativ (als explizite Sollens- oder Könnensbeschreibung) oder terminologisch (als erlernbare Problemlösefähigkeit) zum Einsatz kommt. Je nach Verwendungsart ist auch das Verständnis von Kompetenz und Kompetenzorientierung ein anderes. Blickt man in Literaturdatenbank des Fachinformationssystems Bildung (FIS Bildung) des Deutschen Instituts für Internationale Pädagogische Forschung (Stand 26. März 2018), so erscheinen zum Schlagwort Kompetenz allein 5630 Sammelwerksbeiträge und 16105 Zeitschriftenartikel, die seit 2004 publiziert wurden. Frey (2004, S. 904) fasst dabei das wissenschaftliche Verständnis von Kompetenz wie folgt zusammen:

Besitzt eine Person Kompetenz, so kann sie etwas, ist handlungsfähig und übernimmt für sich und andere Verantwortung. Sie kann so tätig werden, dass sie ein Ziel oder einen Zweck unter Beachtung von Handlungsprinzipien, Werten, Normen und Regeln mit Bezug auf konkrete, die jeweilige Handlungssituation bestimmende Bedingungen zu erreichen vermag. Wer Kompetenz besitzt, ist erfolgreich und vernünftig tätig. Somit kann man Kompetenz als ein Bündel von körperlichen und geistigen Fähigkeiten bezeichnen, die jemand benötigt, um anstehende Aufgaben oder Probleme zielorientiert und verantwortungsvoll zu lösen, die Lösungen zu bewerten und das eigene Repertoire an Handlungsmustern weiterzuentwickeln.

In der psychologisch-pädagogischen Diagnostik taucht der Begriff immer wieder auf „als Gegenbegriff zu generalisierten, kontextunabhängigen kognitiven Leistungskonstrukten [...], wie sie für die Intelligenzforschung und -diagnostik typisch sind“ (Klieme, Maag Merki & Hartig, 2007, S. 6). Ziel solcher Forschungen war häufig die bessere Vorhersagbarkeit der Leistungsfähigkeit einer Person mittels geeigneter Testinhalte und die Erklärung von Leistungsunterschieden in konkreten (zumeist beruflichen) Situationen, für deren Bewältigung Kompetenzen im weitesten Sinne von McClelland (1973) als notwendige Voraussetzungen betrachtet wurden. Wenn in der Pädagogik von Kompetenz gesprochen wird, bezieht sich dies zumeist nicht auf den Erziehenden, sondern auf den Erzieher, der in seiner Erziehungsarbeit Wissen und Können miteinander in Einklang bringen können soll. Zur pädagogischen Handlungskompetenz zählte Elbing (1978, S. 37f.) zum einen den Umfang und die Qualität des handlungsbezogenen Wissens, zum anderen die Art der Wissensverwendung. In ähnlicher Weise wird auch in dieser Studie von dem Grundverständnis ausgegangen, dass eine Kompetenz „anwendungsfähiges Wissen und ganzheitliches Können“ (Klieme & Rakoczy, 2008, S. 222) erfordert, um durch Strategien (Baumert, Stanat & Demmrich, 2002, S. 22) oder Problemlösefähigkeit (Hu, 2008, S. 3f.) bestimmte

² Auch im Englischen wird der Begriff *competence* im Oxford Dictionary ähnlich beschrieben als „the ability to do something successfully or efficiently“ (www.oxforddictionaries.com, Zugriff am 26.08.2013) oder im Oxford Advanced Learner's Dictionary als „a skill that you need in a particular job or for a particular task“ (www.oup.com/elt/oald/, Zugriff am 26.08.2013).

³ Auch eine Ableitung vom lateinischen Verb *competere* (übersetzt mit zusammentreffen) oder vom Substantiv *competentia* (übersetzt mit Zusammentreffen oder Zuständigkeit) ist erklärbar und findet sich in zahlreichen Ansätzen wieder, wenn man davon ausgeht, dass zur Bewältigung einer Situation „die Erfordernisse der Situation mit dem individuellen Konglomerat von Fähigkeiten und Fertigkeiten eines Menschen ‚zusammentreffen‘“ (Wollersheim (1993, S.89)).

kontextgebundene Anforderungssituationen bewältigen zu können. Diesem Anspruch der tatsächlichen Nutzung und Anwendung folgend werden Kompetenzen in der angloamerikanischen Bildungsforschung auch als „realized abilities“ (Connel, Sheridan & Gardner, 2003, S. 142) bezeichnet.

Die zu untersuchende Anforderungssituation stellt in dieser Studie das Beurteilen der Schlagwurfbewegung dar, wofür der Beurteilende in der Lage sein muss, das Wissen über diese sportliche Bewegung so anzuwenden, dass eine zielgerichtete Beobachtung geeigneter Bewegungsmerkmale möglich ist. Dabei ist die Mehrdimensionalität von Kompetenz als „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten [zu betonen], um bestimmte 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“ (Weinert, 2001b, S. 274). Diese Mehrdimensionalität deutet sich in dieser Studie dahingehend an, dass beispielsweise das Fachwissen über die richtige oder auch fehlerhafte Bewegungsausführung sowie die Qualität der Eigenrealisation der zu beurteilenden Bewegung zwei mögliche Variablen darstellen, welche die Kompetenz der Beurteilung sportlicher Bewegungen maßgeblich beeinflussen könnten. In der Psychologie hat sich der Kompetenzbegriff eher fähigkeitsorientiert ausgeprägt⁴ und wird als Steuerungsmechanismus menschlichen Handelns betrachtet, welcher aus einer nichtformulierten Wissensbasis heraus auch das Erzeugen von Handlungen ermöglicht. Hier ist eine deutliche Anlehnung an Chomsky erkennbar, indem Kompetenz sowohl handlungsrelevantes Wissen als auch das Verfügenkönnen über dieses Wissen im Rahmen von Handlungen beinhaltet (Vonken, 2005, S. 30f.).

Zugleich ist festzuhalten, dass der Kompetenzbegriff ein theoretisches Konstrukt ist, mit dessen Hilfe es möglich sein sollte, die zugrundeliegenden Tiefenstrukturen einer gezeigten Leistung (Performanz) zu ermitteln. Die Kompetenz selbst ist dabei nicht beobachtbar, sondern nur das, was Menschen mithilfe ihrer Kompetenz konkret leisten oder wie sie sich in bestimmten Situationen verhalten (Meyer, 2012, S. 147). Somit schlussfolgert Klieme, Avenarius et al. (2007, S. 72ff.), dass Kompetenz als Verbindungsglied zwischen kognitivem Wissen und operativem Können nur leistungsbezogen (d. h. in konkreten Leistungssituationen) erfasst und gemessen werden kann und zugleich so komplex ist, dass allein die Darstellung und Erfassung einzelner isolierter Leistungen zur Kompetenzbestimmung nicht ausreichen. Dementsprechend soll auch in dieser Studie eine leistungsbezogene Kompetenzmessung erfolgen, indem Testpersonen konkrete Bewegungsausführungen eines Schlagwurfs anhand verschiedener Bewegungsmerkmale beobachten und beurteilen müssen und über den Vergleich mit Experteneinschätzungen die Qualität des Beurteilungsprozesses operationalisiert wird (siehe Kapitel 4.5).

Um die empirische Überprüfbarkeit einer Kompetenz zu ermöglichen, muss eine Engführung des Kompetenzbegriffs erfolgen. Daher schlägt Hartig (2008, S. 17) in Anlehnung an Weinert (1999) eine relativ eng gefasste Arbeitsdefinition vor, wonach Kompetenzen als erlernbare, kontextspezifische und kognitive Leistungsdispositionen verstanden werden, die sich nach Klieme, Maag Merki et al. (2007, S. 6) auch als Kenntnisse, Fertigkeiten oder Routinen charakterisieren lassen. Motivationale und affektive Faktoren beispielsweise sollen in den Kompetenzbegriff nicht hineinspielen, denn sonst verliere das Konstrukt an Präzision, Aussagekraft und Stabilität (Hartig, 2008, S. 19). Ähnlich verhält es sich mit der konzeptuellen Abgrenzung und separaten Messung von allgemeinen kognitiven Fähigkeiten, die aufgrund ihrer Kontextunabhängigkeit nicht als Kompetenzen betrachtet werden. Indem möglichst viele

⁴ Beispielsweise „als Fähigkeit zur Kontrolle, als Fähigkeit zur Akkomodation und Assimilation, als Fähigkeit zur Auseinandersetzung mit oder Bewältigung von (neuen) Situationen, als erwiesene Fähigkeit zum Handeln, als Fähigkeit zur selbständigen und eigenverantwortlichen Lebensführung“ (Wollersheim, 1993, S.92).

interindividuell variierende Einflussgrößen aus dem Kompetenzbegriff genommen werden, können Effekte unterschiedlicher individueller Ausgangsbedingungen sowie Entwicklungsprozesse im Bildungsgeschehen differenzierter untersucht werden (Klieme, Maag Merki et al., 2007, S. 7f.). Daher soll in dieser Studie über die Untersuchung der kontextspezifischen Leistungsdispositionen, über deren Bedeutung für die Kompetenzausübung aktuell in den meisten Beiträgen zur Bildungsforschung Konsens herrscht, die Bedeutungen möglicher Einflussfaktoren auf die Beurteilungsleistung festgestellt werden, um möglicherweise Rückschlüsse auf die Erlernbarkeit der Kompetenz zu ziehen.

Es besteht grundsätzliche Einigkeit in der Kompetenzforschung, dass Kompetenzen erlernbar sind. Im Gegensatz zu kognitiven Grundfunktionen seien die für die Bewältigung einer konkreten Situation erforderlichen funktionalen Dispositionen leichter erlernbar, durch äußere Interventionen beeinflussbar und somit prädestiniert für Vermittlung (siehe auch Baumert et al. (2002); Weinert (2001a); Klieme & Leutner (2006, S. 874)). Kompetenzen können bzw. müssen also durch Sammeln von Erfahrungen in entsprechend aufbereiteten Situationen oder Aufgaben erworben werden (Hartig & Klieme, 2006). Wenn das Gefüge möglicher Einflussfaktoren bekannt ist, können entsprechende Lernaufgaben entwickelt werden, die dann auch die spätere Kompetenzausübung in realen Situationen positiv beeinflussen können. Durch die Dynamik dieser Wissens- und Handlungsbasis ist eine Anpassung an zukünftigen Anforderungen gegeben, wenngleich nicht von einer kurzfristigen Trainierbarkeit der Kompetenzen auszugehen ist. Erst der autonome Gebrauch vorhandener Fähigkeiten und Fertigkeiten, Motivationen und Volitionen stelle die Anschluss- und Zukunftsfähigkeit von Kompetenzen sicher.

Frey (2006, S. 32) unterscheidet dabei drei Entwicklungsebenen, in denen in jedem Lebensalter durch Bewältigung von Aufgaben oder Problemen neue Wissens- und Könnensbasen entwickelt werden. Über Tätigkeiten als organisierende zentrale Aktivität werden grundlegende Fertigkeiten erlangt, die sich zu Fähigkeiten (als hoch entwickelte Fertigkeiten) und auf der obersten Ebene zu Kompetenzen (als hoch entwickelte Fähigkeiten) entwickeln können (siehe Abb. 4). Während die Fertigkeit ein konkretes und inhaltlich bestimmbares Können darstellt, das weitgehend automatisch auch unter Ausschaltung des Bewusstseins abgerufen werden kann, stellt die Fähigkeit schon ein theoretisches Konzept dar, welches alle tiefer liegenden psychischen und physischen Fertigkeiten planvoll bündelt. Auf der Kompetenzebene werden diese Fähigkeitskonzepte als theoretische Klasse zu einer Einheit zusammengefasst. Es ist also davon auszugehen, dass auch Probandinnen und Probanden dieser Studie, die schon häufig Beurteilungstätigkeiten wahrgenommen haben, ein höheres Leistungsniveau beim Beurteilen der Schlagwurfbewegung erreichen als die unerfahreneren Testpersonen. Dementsprechend ist anzunehmen, dass auch die Vorerfahrung im Beurteilen sportlicher Bewegungen ein möglicher Einflussfaktor auf die Beurteilungsleistung ist (siehe Kapitel 4.6).

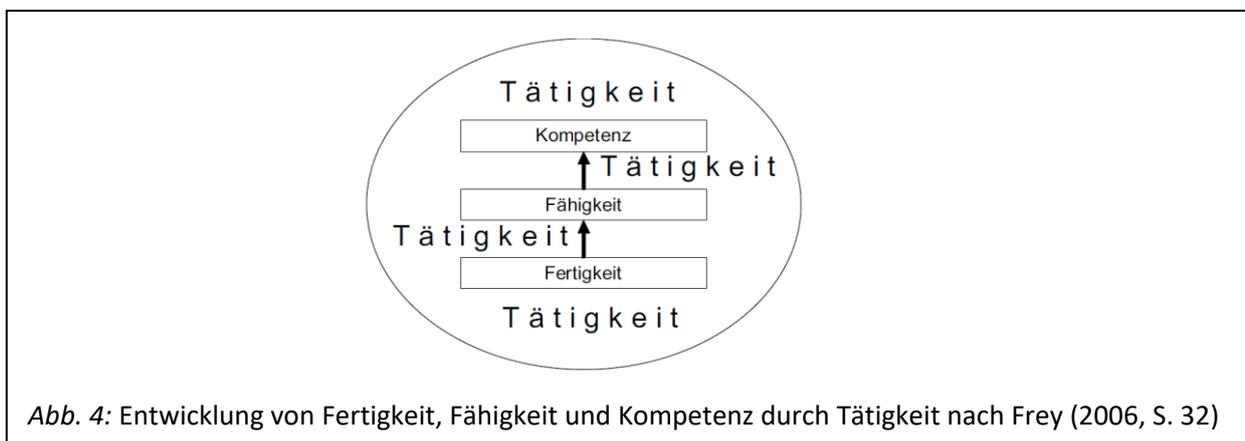


Abb. 4: Entwicklung von Fertigkeit, Fähigkeit und Kompetenz durch Tätigkeit nach Frey (2006, S. 32)

Eine Kompetenzbeschreibung sollte also immer von der Frage ‚Kompetent wofür?‘ ausgehen und grenzt sich in diesem Sinne von der Intelligenzforschung ab, die nach situationsunabhängigen, generalisierbaren kognitiven Leistungskonstrukten sucht (Klieme, Maag Merki et al., 2007, S. 6). Kompetent handeln bedeute vielmehr, spezifische Situationsanforderungen in Lernfeldern oder fachbezogenen Leistungsbereichen (werden in der Expertiseforschung als sog. Domänen bezeichnet) bewältigen zu können. Klieme und Leutner (2006, S. 879f.) stellen somit den Aspekt der **Kontextabhängigkeit** in den Mittelpunkt, wobei mit Kontext konkrete Situationen und Anforderungen gemeint sind, auf welche sich ein spezifisches Kompetenzkonstrukt bezieht.

Entscheidend für die Qualität von Bildungsprozessen ist dabei die passende Auswahl des Kontextes. Problematisch wird es immer dann, wenn Leistungsdispositionen für einen breiten Bereich von Situationen festgelegt werden, so dass keine Kontextspezifität mehr vorliegt, weil dann das Konstrukt Kompetenz zu inhaltlicher Beliebigkeit zu verkommen droht. Dies ist häufig bei fachübergreifenden oder sog. Schlüsselkompetenzen (z. B. Problemlösefähigkeit) der Fall, wenn die charakteristische Kontextabhängigkeit nicht gegeben ist (Klieme, Maag Merki et al., 2007, S. 15). Ebenso kann ein zu enger Kontext den Kompetenzbegriff entwerten, wenn nämlich isolierte Fertigkeiten oder Sachwissen unnötig als Kompetenzen etikettiert werden. Darum müsse vor der Kompetenzformulierung grundlegend überlegt werden, auf welchen Bereich von Situationen und Anforderungen sich ein jeweiliges Kompetenzkonstrukt beziehen soll. Ein Gradmesser für die Spezifität des Kontextes könne die Menge realer (d. h. außerhalb des Bildungsprozesses liegender) Situationen sein, in denen hinreichend ähnliche Anforderungen zu bewältigen seien (Klieme, Maag Merki et al. (2007, S. 8), Hartig (2008, S. 18–20)).

Es ist also zwingend erforderlich, das Kompetenzkonstrukt hinreichend konkret, aber nicht zu eng zu fassen, so dass die Kompetenzerwartungen einerseits noch überprüfbar bleiben und andererseits aber auch noch Raum für individuelle Lösungs- und Lernwege gelassen wird. Dies gilt es vor allem für die Evaluation von Bildungsprozessen, in denen immer spezifische Fähigkeiten für spezifische Kontexte vermittelt werden, zu berücksichtigen: Wenn Kompetenzklassifikationen zu global sind, kann der eigentlich intendierte, spezifische Kompetenzgewinn nicht mehr eindeutig abgebildet und überprüft werden (Hartig, 2008, S. 21f.). Durch die Kontextspezifität können jedenfalls spezifische Kompetenzen formuliert werden, aus welchen wiederum „Kriterien zur Evaluation in *Passung zu den intendierten Zielen der zu evaluierenden Maßnahmen*“ (Klieme, Maag Merki et al., 2007, S. 7) abgeleitet werden können. Als kontextspezifische Situationen kommen daher in dieser Studie gefilmte Bewegungsausführungen von Schülerinnen und Schülern zum Einsatz, mit denen Lehrkräfte auch im realen Unterrichtsalltag häufig konfrontiert werden.

Zusammenfassend kann festgehalten werden, dass das Beurteilen sportlicher Bewegungen durchaus den theoretischen Anforderungen an eine Kompetenz gerecht wird. Die Bewegungsbeurteilung stellt ohne Zweifel eine häufig wiederkehrende Situation im Trainings- und Unterrichtsalltag von Trainern und Sportlehrkräften dar, bei der verschiedene kognitive Leistungsdispositionen synchron angewendet werden, um eine sportliche Bewegung wie den leichtathletischen Schlagwurf beobachten und im Vergleich zur idealen Bewegungsausführung beurteilen zu können. Die Anwendungssituation erweitert sich noch, wenn man bedenkt, dass die Bewegungsbeurteilung in eine Rückmeldung über die Qualität der Bewegungsausführung zum Zwecke der Technikoptimierung mündet oder bei der Benotung einer Bewegungsdemonstration eine hohe Relevanz besitzt. Die zu bewältigende Herausforderung dieser Studie besteht dabei in der Messbarkeit der Beurteilungsleistung und der Identifikation möglicher Einflussfaktoren. Die theoretische Auseinandersetzung mit dem Kompetenzbegriff hat jedenfalls schon wichtige grundlegende Erkenntnisse geliefert, die bei der Testkonstruktion beachtet werden müssen.

2.1.3 Bewegungsbeurteilung als diagnostische Kompetenz

Um den Stellenwert zu klären, den die Bewegungsbeurteilung bei der professionellen Berufsausübung einer Sportlehrkraft einnimmt, erfolgt eine theoretische Auseinandersetzung mit dem allgemeinen Begriff der diagnostischen Kompetenz. Nachdem das Kompetenzverständnis dieser Studie schon in Kapitel 2.1.2 entwickelt wurde, wird nun anhand verschiedener Modelle aus der Bildungsforschung (siehe Kapitel 2.1.3.1) sowie des Kompetenzmodells der Kultusministerkonferenz (siehe Kapitel 2.1.3.2) untersucht, in welcher Form das Beurteilen eine Rolle spielt und wie es in der universitären Lehrerbildung verankert ist. In einem zweiten Schritt wird überlegt, wie diese allgemeinen Kompetenzformulierungen auf die Profession einer Sportlehrkraft übertragen werden kann, indem unter dem Begriff der fachdidaktischen Diagnosekompetenz die speziellen Anforderungen beim Beurteilen sportlicher Bewegungen diskutiert werden (siehe Kapitel 2.1.3.3).

2.1.3.1 Diagnostische Kompetenz in der Bildungsforschung

Im Handwörterbuch Pädagogische Psychologie wird diagnostische Kompetenz im Sinne von Urteilsgenauigkeit ganz allgemein und unabhängig von einer bestimmten Zielgruppe beschrieben als „die Fähigkeit eines Urteilers, Personen zutreffend zu beurteilen“ (Schrader, 2010, S. 95). Artelt und Gräsel (2009, S. 157) konkretisiert für den schulischen Kontext, dass sich die Urteilsfähigkeit von Lehrkräften zum einen auf die lern- und leistungsrelevanten Merkmale ihrer Schülerinnen und Schüler bezieht, zum anderen aber auch die Fähigkeit einschließt, Lern- und Aufgabenanforderungen hinsichtlich ihrer Schwierigkeit so adäquat einzuschätzen, dass sie von einem Großteil einer konkreten Lerngruppe gelöst werden können. Wie hoch dabei die personenbezogene Diagnosefähigkeit einer Lehrkraft ist, kann am Grad der Übereinstimmung zwischen dem mehr oder weniger genauen Lehrerurteil und der tatsächlichen Merkmalsausprägung (d. h. zumeist der erzielten Leistung) eines Schülers festgemacht werden (Helmke, Hosenfeld & Schrader, 2004, S. 120). Mit Brunner, Anders, Hachfeld und Krauss (2011, S. 215) ist also festzuhalten, dass zum einen bei der Vergabe von Noten, zum anderen für den Lernfortschritt der Schülerinnen und Schüler die diagnostische Fähigkeit von Lehrkräften eine Rolle spielt.

Für den Sport stellte Maraun (1981) schon Anfang der 1980er Jahre fest, dass „Analysieren ein Alltagsproblem unterrichtlichen Handelns ist“ (S. 11). Der Fokus ihrer Forschung war auf die Analyse von Handlungsgewohnheiten einer Sportlehrkraft sowie die Analyse von Unterricht gerichtet, welche immer dann distanziert hinterfragt werden müssen, wenn Unterricht als nicht erfolgreich bewertet wird. In der aktuellen Forschung steht hingegen mehr die Schülerleistung im Mittelpunkt, welche von der Lehrkraft im Unterricht richtig eingeschätzt und optimal entwickelt werden soll. Darum wird in dieser Studie die Fähigkeit zur Beurteilung von Bewegungsmerkmalen untersucht, welche dann wiederum als Grundlage für weitere fachdidaktische Prozesse dient. Auf der Grundlage von diagnostischen Prozessen ist es Lehrkräften dann möglich, in der Unterrichtsgestaltung und gezielten individuellen Förderung (beispielsweise im Hinblick auf Aufgabenwahl, Rückmeldung oder Bewertung) pädagogisch verantwortlich zu entscheiden und zu handeln (Artelt & Gräsel, 2009, S. 157). Dabei lassen sich formelle Diagnosen, die explizit theoriegeleitet und standardisiert mit wissenschaftlich geprüften Methoden erstellt werden, von informellen, implizit subjektiven Urteilen, Einschätzungen und Erwartungen unterscheiden, die zumeist spontan und beiläufig sowie unsystematisch im Rahmen des alltäglichen Unterrichtshandelns gewonnen werden (Schrader, 2010, S. 96). Beide Herangehensweisen haben in der Unterrichtspraxis ihre Berechtigung und ergänzen sich gegenseitig.

Neben allen theoretischen Überlegungen über deklarative Wissenskomponenten soll aber nicht vergessen werden, dass diagnostische Kompetenz auch prozedurales (methodisches) Handlungswissen beinhaltet. In diesem Sinne bescheinigt Helmke (2014) einer Lehrkraft erst dann diagnostische Expertise, wenn sie Methoden zur Einschätzung von Schülerleistungen sowie zur Selbstdiagnose nicht nur kennt, sondern auch anwenden kann, sich dabei (als konzeptuelles Wissen verstandener) typischer Urteilstendenzen und -fehler bewusst ist und darüber hinaus „ein hohes Niveau an zutreffender Orientiertheit“ (S. 85) besitzt.

Schon in den 1990er Jahren betonen Weinert, Schrader und Helmke (1990), dass die diagnostische Kompetenz (neben der Klassenführungscompetenz, didaktischen Kompetenz und fachwissenschaftlichen Kompetenz) eine von vier Schlüsselkompetenzen von Lehrkräften darstellt. Im Kompetenzmodell, das der COACTIV-Studie zugrunde liegt, wird hierarchisch in Kompetenzaspekte, Kompetenzbereiche und Kompetenzfacetten unterteilt. Betrachtet man die diagnostische Kompetenz, so ist diese eine Kompetenzfacette des fachdidaktischen Wissens, welches wiederum einen von fünf Kompetenzbereichen des Professionswissens darstellt (siehe Abb. 5).

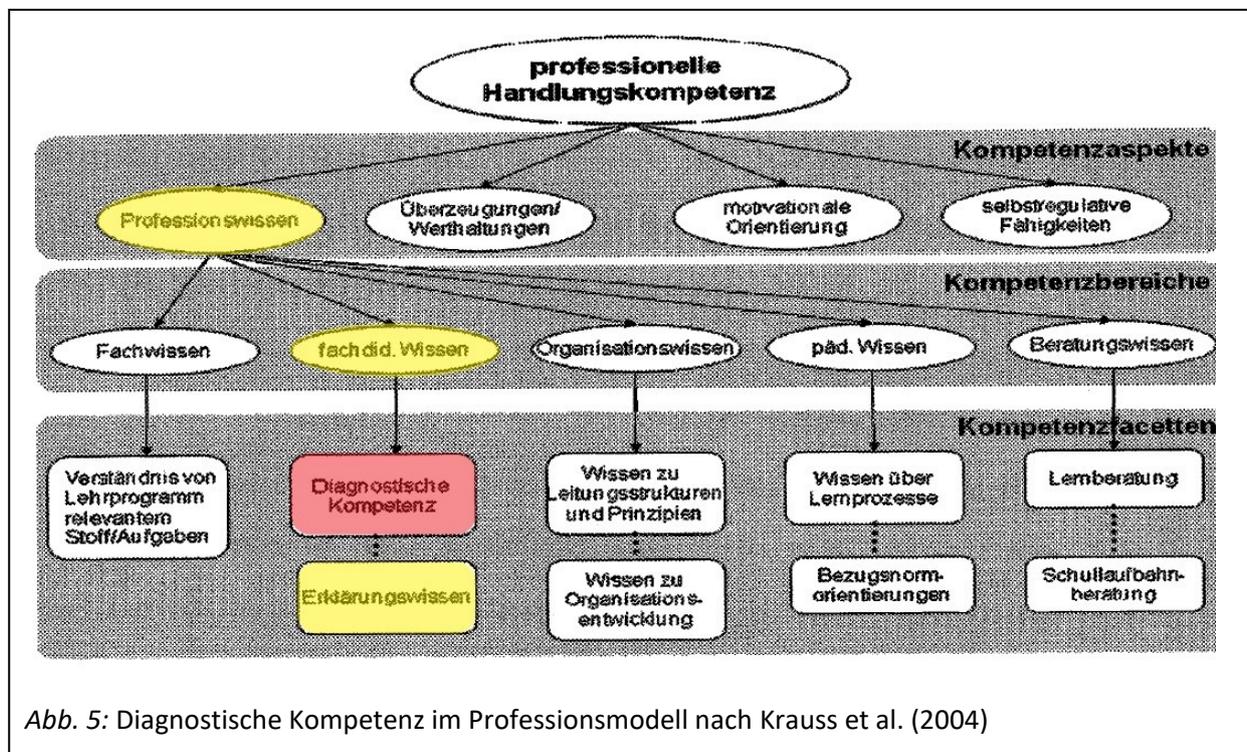


Abb. 5: Diagnostische Kompetenz im Professionsmodell nach Krauss et al. (2004)

Vom gleichen Kompetenzmodell ausgehend, aber etwas anders akzentuiert, sprechen Brunner et al. (2011, S. 216) bewusst nicht von diagnostischer Kompetenz, sondern von diagnostischen Fähigkeiten, um anzuzeigen, dass die untersuchten Fähigkeiten nur einige zentrale Indikatoren einer inhaltlich breiter gefassten, multidimensionalen Kompetenzfacette darstellen. Dabei erstrecken sich diese Indikatoren über die beiden Kompetenzbereiche pädagogisch-psychologisches und fachdidaktisches Wissen. Ersterem ist das Wissen um die Prüfung und Bewertung von Schülerleistungen zuzuordnen, während letzteres sowohl Wissen über fachspezifische Instruktionsstrategien und Wissen über das Potenzial fachspezifischer Aufgaben wie auch Wissen über fachbezogene Schülerkognitionen umfasst (siehe Abb. 6).

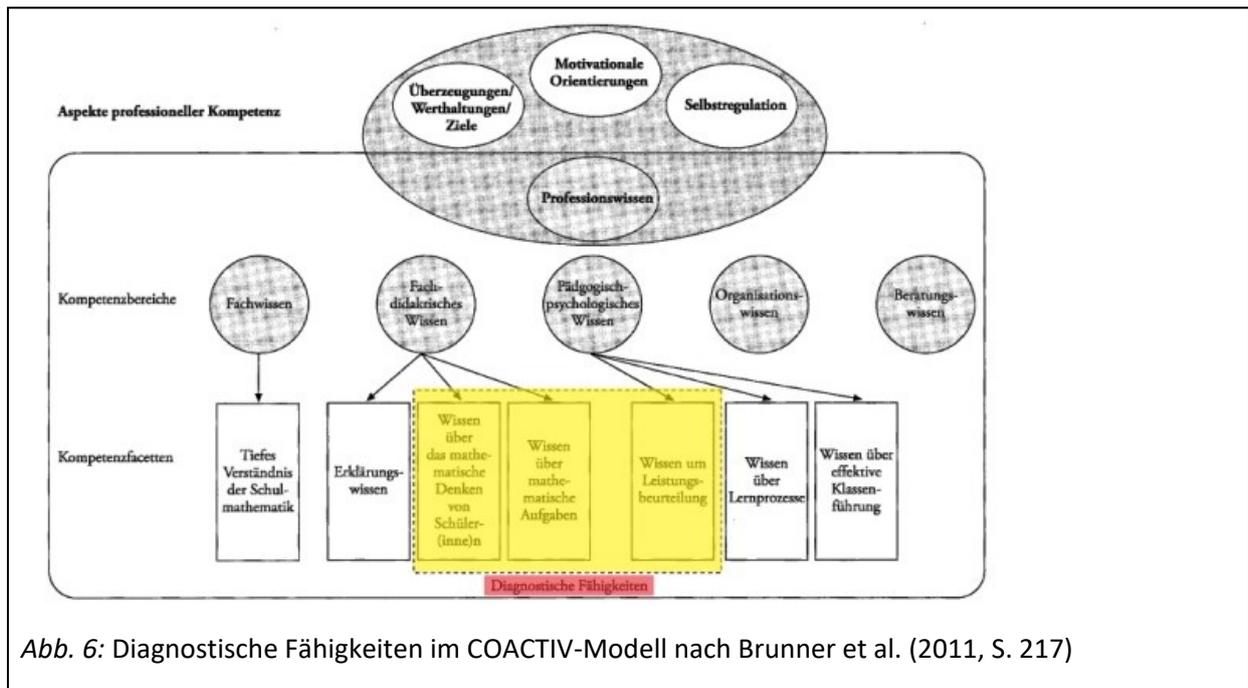


Abb. 6: Diagnostische Fähigkeiten im COACTIV-Modell nach Brunner et al. (2011, S. 217)

Baumert und Kunter (2011, S. 37–39) führt die drei in Abb. 6 genannten diagnostischen Fähigkeiten näher aus. Das fachdidaktische Aufgabenwissen kann als „Wissen über das didaktische und diagnostische Potenzial, die kognitiven Anforderungen und impliziten Wissensvoraussetzungen von Aufgaben, ihre didaktische Sequenzierung und die langfristige curriculare Anordnung von Stoffen“ (S. 37) verstanden werden, wovon das Wissen über Schülervorstellungen, deren Fehlkonzeptionen, typische Fehler und Strategien, sowie die Diagnostik von Schülerwissen und Verständnisprozessen abgegrenzt werden kann. Zum allgemeinen pädagogisch-psychologischen Wissen um Leistungsbeurteilung gehören hingegen fachübergreifende Prinzipien des Diagnostizierens, Prüfens und Bewertens, die mit den Stichworten Lernen und Leisten, Prozessdiagnostik, Rückmeldungen sowie summatives Prüfen und Bewerten beschrieben werden.

Auch im Kompetenzmodell von Oser (2001, S. 230), das in zwölf verschiedenen Standardgruppen insgesamt 88 Standards beschreibt, findet sich die Ausbildung von diagnostischer Kompetenz in folgenden zwei Standardgruppen wieder:

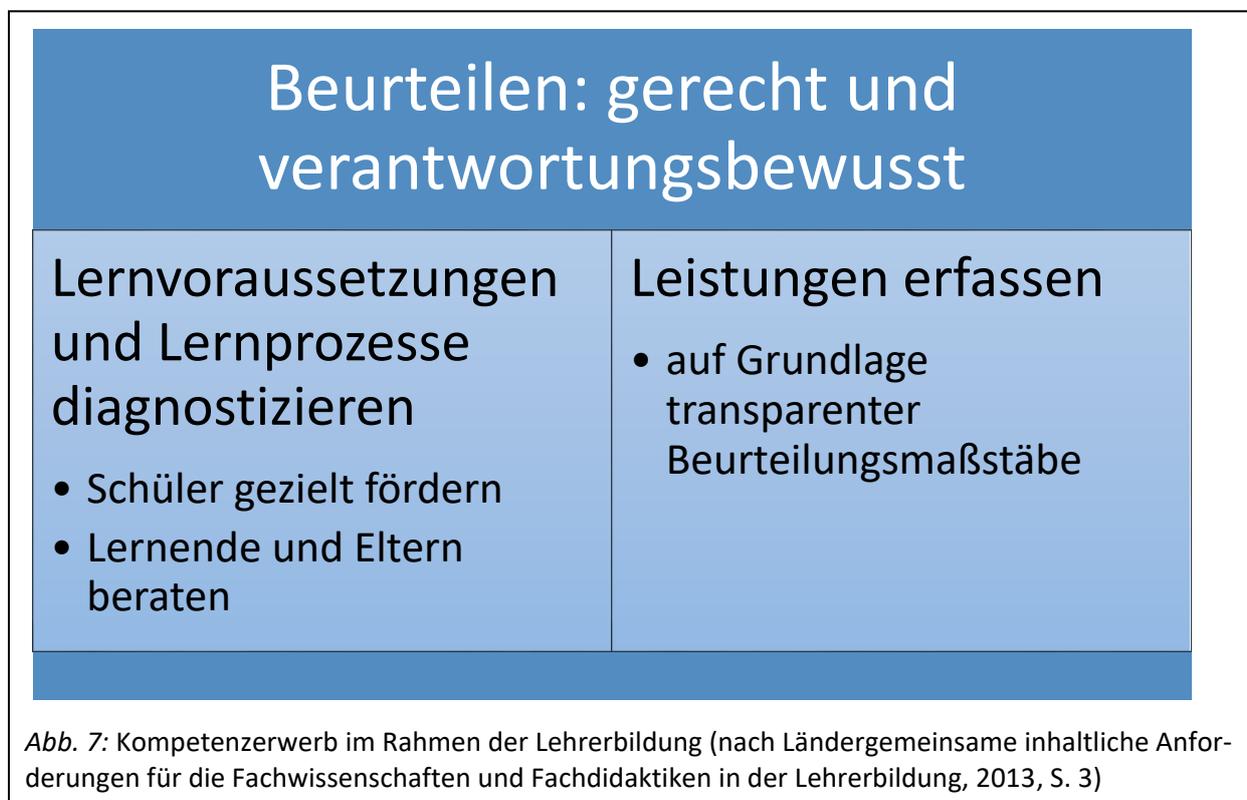
1. Schüler unterstützende Beobachtung (Diagnose) und Schüler unterstützendes Handeln (Intervention): spezifische Lernschwierigkeiten diagnostizieren und beheben können
2. Leistungsmessung: den Fortschritt der Leistungen nach unterschiedlichen Kriterien und mit unterschiedlichen Instrumenten messen (Oser, 2001, S. 232 und S. 237)

2.1.3.2 Beurteilungskompetenz im KMK-Modell

In Kompetenzmodellen wird Kompetenz häufig als „latente (kognitive) Fähigkeit“ (Lehmann-Grube & Nickolaus, 2009, S. 63) verstanden, die nicht direkt messbar ist. Somit kommt der Operationalisierung der im Beruf zu bewältigenden Anforderungen eine entscheidende Bedeutung zu, anhand deren Bewältigungsgrad (Performanz) Rückschlüsse auf die tatsächliche Kompetenz im Sinne der kognitiven Leistung gezogen werden können. Dies geschieht in der pädagogisch-empirischen Forschung häufig,

indem die Grundmuster von Lehrerkompetenz als Variablenausprägung erfolgreichen Handelns bestimmt werden. Die Schwierigkeit besteht vor allem in den sich ständig ändernden Anforderungen an den Lehrerberuf, wodurch die Erfassung elementarer Fähigkeiten und Fertigkeiten niemals vollständig sein kann (Cramer, 2012, S. 36).

In einem Beschluss der Kultusministerkonferenz vom 16.12.2004 (Standards für die Lehrerbildung: Bildungswissenschaft, 2004) wird das Beurteilen als einer von vier Kompetenzbereichen (neben dem Unterrichten, Erziehen und Innovieren) genannt, die für die berufliche Ausbildung (Lehramtsstudiengänge einschließlich praktischer Ausbildungsteile und Vorbereitungsdienst) und den Berufsalltag von besonderer Bedeutung sind und regelmäßig evaluiert werden sollten. Geht man von den Begriffen der Diagnostik aus, so kann man die vier Kompetenzbereiche als Skalen bezeichnen, die in sog. Subskalen näher spezifiziert werden, aus denen wiederum Standards, Aufgaben und Items für die Kompetenzdiagnostik abgeleitet werden können (Hartig & Jude, 2007, S. 18). Die Standards für die Lehrerbildung (2004, S. 4) betonen aber zugleich, dass die Formulierung dieser allgemeinen bildungswissenschaftlichen Kompetenzen und Standards sich an fachlichen Inhalten vollziehen muss und somit im Rahmen fachdidaktischer Forschung jeweils konkretisiert werden muss. Eine Schwerpunktsetzung erfolgte schon dahingehend, dass die Standards für theoretische und praktische Ausbildungsabschnitte getrennt voneinander formuliert wurden, ohne dies jedoch als gegenseitige Abgrenzung verstehen zu wollen. Für den Kompetenzbereich Beurteilen werden das Diagnostizieren von Lernvoraussetzungen und Lernprozessen sowie das Erfassen von Leistungen als Kernkompetenzen genannt (siehe Abb. 7), wobei die Leistungserfassung auf der Grundlage transparenter Beurteilungsmaßstäbe den zentralen Gegenstand dieser Studie darstellt.



Betrachtet man die grundlegenden Kompetenzen genauer, die im Rahmen des Lehramtsstudiums entwickelt werden sollten, so geht es darum, Fachwissen und fachdidaktische Kenntnisse so zu erwerben, dass sie anschlussfähig sind für spätere berufliche Anforderungen. Dazu gehört die Fähigkeit, anforderungsgerechte Leistungsbeurteilungen vornehmen und differenziert Einfluss auf den Lernerfolg der Schüler nehmen zu können. Mit dem Abschluss des Studiums sollen Studierende befähigt sein, eine „fach- bzw. fachrichtungsspezifische Leistungsbeurteilung [zu] beherrschen“ (Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung, 2013, S. 4)

In den jeweiligen Fachprofilen werden aus den zu erreichenden allgemeinen Kompetenzen konkrete inhaltliche Schwerpunkte (siehe Abb. 8) abgeleitet. So wird zunächst ein fachspezifisches Kompetenzprofil für Sport entworfen mit dem Anspruch, Themenbereiche zu identifizieren, die zukünftige Sportlehrkräfte auf die schulische Tätigkeit vorbereiten sollen (Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung, 2013, S. 48). Dabei wird neben der verantwortungsvollen Leistungsbeurteilung auch das Analysieren von Bewegungslernsituationen explizit als fachdidaktische Kernkompetenzen genannt, aus welchen wiederum Studieninhalte wie Bewegungskompetenz und sportliches Können oder Bewegungstheorie und sportliches Handeln abgeleitet werden. Für den Bereich Unterrichten und Vermitteln von Bewegung, Spiel und Sport wird dabei das Lehren und Lernen von Bewegungen als Konkretisierung genannt, für welche auch das Beurteilen von Bewegungen Relevanz hat. Im Bereich Bewegungskompetenz und sportliches Können soll neben dem Sich-Bewegen auch das Wahrnehmen von Bewegungen als Grundlage des Bewegungskönnens geschult werden. Somit kann das Beurteilen durchaus als immer wiederkehrender Prozess in verschiedenen Studieninhaltsbereichen verstanden werden.

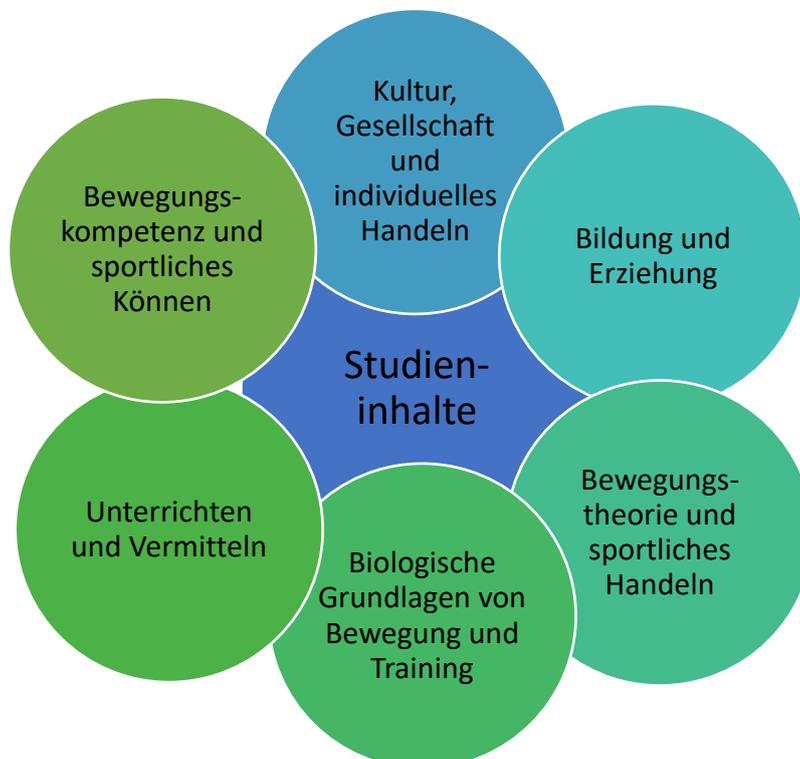


Abb. 8: Aus den Kompetenzen abgeleitete Studieninhalte nach dem KMK-Modell (Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung, 2013, S. 48)

Der Einteilung in die vier Kompetenzbereiche Unterrichten, Erziehen, Beurteilen und Innovieren liegt jedoch für Girmes (2006, S. 17–19) eine theoretisch unzureichend begründete Argumentation zugrunde. Zwar wird dadurch betont, dass nicht mehr Inhalte von wissenschaftlichen Disziplinen im Vordergrund stehen, sondern professionelle Tätigkeiten, diese können allerdings nicht als Standards für die Lehrerbildung dienen. Unterrichten und Erziehen stellen keine voneinander klar abgrenzbaren Kompetenzen dar, sondern sind lediglich zwei verschiedene Ausprägungen von lehrender Kommunikation auf Sachebene und beziehen sich somit auf curriculare, methodische, institutionelle und kulturelle Kompetenz. Innovieren ist ebenfalls keine eigenständige Kompetenz, weil sie sich immer auf Bereiche wie Curriculum, Methoden, Institution, Lernumgebung oder Bewertungsprozesse bezieht. Darum sei es schlüssiger, Innovieren als Bestandteil der jeweiligen Kompetenzen zu betrachten. Dass Diagnoseprozesse nicht genannt werden, erscheint angesichts der Wichtigkeit in der schulischen Realität ein grundlegender Mangel.

Terhart (2007, S. 48f.) konkretisiert die Aufgabenfelder des KMK-Modells und betont dabei, dass die zentrale Aufgabe eines Lehrers immer noch das Unterrichten darstelle, weshalb sich die Qualität eines Lehrers im Kern vor allem an der Qualität seines Unterrichts bemisst und auch die anderen Aufgabenfelder auf das Unterrichten bezogen bleiben. In der Kategorie Beurteilen hebt er die Diagnose von Lernständen sowie den Einsatz unterschiedlicher Formen der Erfassung, Dokumentation und Rückmeldung von Schülerleistungen hervor.

Zwar taucht das Beurteilen in weiteren Lehrerkompetenzmodellen auf (bei Oser (2001) beispielsweise im Begriff der Leistungsmessung), diese sind aber für den weiteren Fortgang dieser Studie nicht hilfreich, da das Beurteilen sehr stark auf kognitive Prozesse von Schülerinnen und Schülern bezogen ist (beispielsweise beim heuristischen Rahmenmodell der Handlungskompetenz nach Baumert & Kunter (2006)) und das Beurteilen sportlicher Bewegungen hier eine sehr kontextspezifische Kompetenz darstellt, die speziell von Sportlehrkräften beherrscht werden muss.

2.1.3.3 Fachdidaktische Diagnosekompetenz im Sport

Wenn man von diagnostischer Kompetenz einer Sportlehrkraft spricht, so sind zwar die in Kapitel 2.1.3.1 genannte bildungswissenschaftlichen Aspekte als Orientierung brauchbar, müssen aber im Hinblick auf die Anforderungen der spezifischen Unterrichtsanforderungen im Schulfach Sport modifiziert werden. Grundsätzlich kommt es im Sportunterricht permanent vor, dass Bewegungen beurteilt werden müssen, sei es im Sinne von Bewegungsbeschreibungen, um den Schülern eine möglichst nachvollziehbare Bewegungsvorstellung zu vermitteln, sei es im Sinne von Bewegungskorrekturen, um den Lernenden ihre Defizite bzw. Fehler bei der Bewegungsausführung bewusst zu machen. Baumann (1986, S. 17) spricht vom „diagnostischen Blick“, den ein künftiger Sportlehrer für sportmotorische Handlungsvollzüge haben oder erwerben soll, um durch Beobachtung Fehler analysieren zu können. Ebenso gehört zur Diagnosekompetenz einer Sportlehrkraft die Fähigkeit, didaktisch reduzierte Bewegungsanweisungen in Form von Lernschritten, Zusatzinformationen oder Korrekturhinweisen zu geben (Hartmann, 1999, S. 106f.). Aschebrock und Pack (2004, S. 28) postulieren im Jahr 2004 sogar, dass im Sportunterricht Übungssituationen für das Bewegungssehen der Schüler geschaffen werden sollen und ihre Beobachtungsergebnisse in das jeweils eigene Bewegungshandeln einfließen, was mittlerweile in den aktuellen kompetenzorientierten Lehrplänen (z. B. in Bayern oder Baden-Württemberg) verankert wurde. Dazu müsse die Sportlehrkraft nicht nur selbst in der Lage sein, Bewegungen zielgerichtet beobachten zu können, sondern auch das Wissen darüber besitzen, wie man diese Handlungskompetenz

bei Schülern anbahnen und sie vom bloßen Hinsehen zum Beobachten anleiten kann (Baumann, 1986, S. 15–17). In dieser Studie steht jedoch ausschließlich die Beobachtung und Beurteilung sportlicher Bewegungen durch die Lehrkraft im Mittelpunkt, welche als Voraussetzung für weitere didaktische Entscheidungen dienen.

Laut Meinel, Schnabel und Krug (2015, S. 377) liegt jedoch noch kein fundiertes Theoriekonzept zur Beobachtung, Analyse und Beurteilung von Bewegungen im Sport vor. Daher wird im Rahmen dieser Arbeit der Begriff der **fachdidaktischen Diagnosekompetenz** als ein neues Konstrukt verwendet, welches die Fähigkeit einer Lehrperson im Sport beschreibt, eine sportliche Bewegung (Schülerleistung) zu analysieren und im Falle einer Abweichung von der altersgemäß adäquaten Bewegungsausführung mit fachdidaktisch begründeten Instruktionsstrategien und Aufgaben zu reagieren, um den Bewegungslernprozess der Schüler zu optimieren.

Dies sei nochmals an einem Beispiel verdeutlicht. Ein grundlegender Unterschied zu kognitiv orientierten Fächern wie Mathematik besteht darin, dass es im Sportunterricht vor allem um das Erlernen und Ausführen von Bewegungen geht. In der Sportart Leichtathletik stellt beispielsweise das schnelle Sprinten eine Schülerleistung dar, die in jeder Jahrgangsstufe geschult und bewertet wird und auch eine wichtige Voraussetzung für andere Disziplinen wie den Weitsprung darstellt. Aus fachdidaktischer Sicht bestünden die diagnostischen Fähigkeiten der Sportlehrkraft folglich darin, altersgemäße Kriterien der leichtathletischen Sprinttechnik theoretisch zu kennen, die Bewegungsausführung von Schülerinnen und Schülern diesbezüglich zu beurteilen und Wissen über leichtathletische Übungen zu besitzen, mit denen man bestimmte Bewegungsausführungen provozieren oder auftretende Fehlerbilder verbessern kann. Im konkreten Fall des leichtathletischen Sprints würde dies bedeuten, dass die Sportlehrkraft das Potenzial der fachspezifischen Aufgaben des Sprint- und Sprung-ABCs kennt, welches darin besteht, dass zentrale leichtathletische Grundfertigkeiten für spätere Sprint- und Sprungdisziplinen geschult werden. Indem die Lehrkraft über fachspezifische Instruktionsstrategien verfügt, kann sie die Übungen fachgerecht anleiten und dem Alter und Niveau der Schüler entsprechend anpassen und verbalisieren. Wenn die Lehrkraft außerdem in der Lage ist, aufgrund eines zuverlässigen diagnostischen Urteils sowohl die kognitiven als auch die motorischen Leistungsvoraussetzungen der Schüler adäquat einzuschätzen, sind die Bedingungen für einen hohen Lernerfolg gegeben.

Die Verwendung des Begriffs **Fehler oder Fehleranalyse** ist dabei nicht unproblematisch. Zum einen weisen Bremer und Sperle (1984, S. 6) darauf hin, dass der Begriff gesellschaftlich negativ gefärbt ist, zugleich aber das Begehen von Fehlern als notwendige Voraussetzung für eine Weiterentwicklung in allen Lebensbereichen (und insbesondere beim motorischen Lernen) gesehen werden kann. Im Sinne des differenziellen Lernens gibt es im sportlichen Lernprozess streng genommen keine Fehler, denn Abweichungen vom gewohnten Bewegungsablauf können dem Herausfinden der potentiellen Streugrenzen einer Bewegungsausführung dienen (Schöllhorn, Beckmann, Janssen & Michelbrink, 2009, S. 38). Und Wolters (2002, S. 17ff.) weist darauf hin, dass überhaupt erst einmal exakt geklärt werden muss, was als Fehler gelten soll, bevor man sich über die Korrektur eines solchen Gedanken machen kann. In dieser Arbeit wird der Begriff Fehler einheitlich als Differenz zwischen Ist- und Sollwert verstanden, unabhängig der Tatsache, dass es auch noch differenziertere Begriffsunterscheidungen (z. B. Mängel oder Abweichung) gibt. Die Begriffsdiskussion soll im Rahmen dieser Arbeit aber nicht weiter behandelt werden.

2.1.3.3.1 Morphologische Betrachtungsweise als Mittel der Bewegungsbeobachtung

Die Bewegungsbeobachtung und Bewegungsbeurteilung dieser Studie basiert auf der morphologischen Betrachtungsweise von Bewegungen. Der Begriff begründet sich nach Meinel (1960) und beschreibt die Fähigkeit einer Sportlehrkraft, „sportliche Bewegungen in ihrem aktuellen Vollzug, noch bevor sie einer weiteren analytischen Betrachtung unterzogen werden (können), einzuschätzen“ (Meinel et al., 2015, S. 377). Die zu beobachtende Bewegung ist also vornehmlich dem Auge direkt zugänglich, wird aber zugleich über den kinästhetischen Analysator miterlebt und von innen her wahrgenommen, wodurch sich eine große Praxisnähe dieser erscheinungswissenschaftlichen Betrachtungsweise ergibt. Durch die Kenntnis von objektiven Faktoren und Gesetzen, die der jeweiligen Bewegungsform zugrunde liegen, soll es der Sportlehrkraft möglich sein, sich durch Vergleich und Abstraktion bestimmte Merkmale und Eigenschaften einer Bewegung zu erschließen, Bewegungsverwandtschaften festzustellen, Zusammenhänge und Beziehungen zu erkennen und letztendlich gewisse Verallgemeinerungen zu treffen (Hartmann, 1999, S. 108).

Diese mit dem Auge beobachtbaren Merkmale als Ausdruck der Bewegungskoordination bezeichnet Neumaier (1988, S. 210f.) als **Beobachtungseinheiten**, welche mittels phänomenologisch-analytischer Betrachtung durch die Auflösung einer Bewegung ermittelt werden können und dem Beobachter ein Höchstmaß an benötigter Information vermitteln. Dabei können nach Baumann (1986, S. 50ff.) auf phänomenaler Ebene zwei Herangehensweisen unterschieden werden, um die Qualität einer Bewegungsbeobachtung zu erklären (siehe Abb. 9). Gestalttheoretisch wird davon ausgegangen, dass die Bewegung (auch als Bewegungsgestalt bezeichnet) als Ganzheit wahrgenommen und ihr Verlauf im Sinne einer Ortsveränderung einer Figur und ihrer Teile im Zusammenhang beschrieben wird.⁵ Informationstheoretisch kann eine Bewegung auch als Superzeichen verstanden werden, welches über den Vorgang der sog. Superierung durch Komplexbildung derart vereinfacht wird, dass raumzeitlich geordnete Bewegungsmuster (auch als Signalkomplexe bezeichnet) zusammengefasst oder unterschiedliche viele Teilinformationen zu größeren Einheiten (sog. chunks) verknüpft werden.⁶

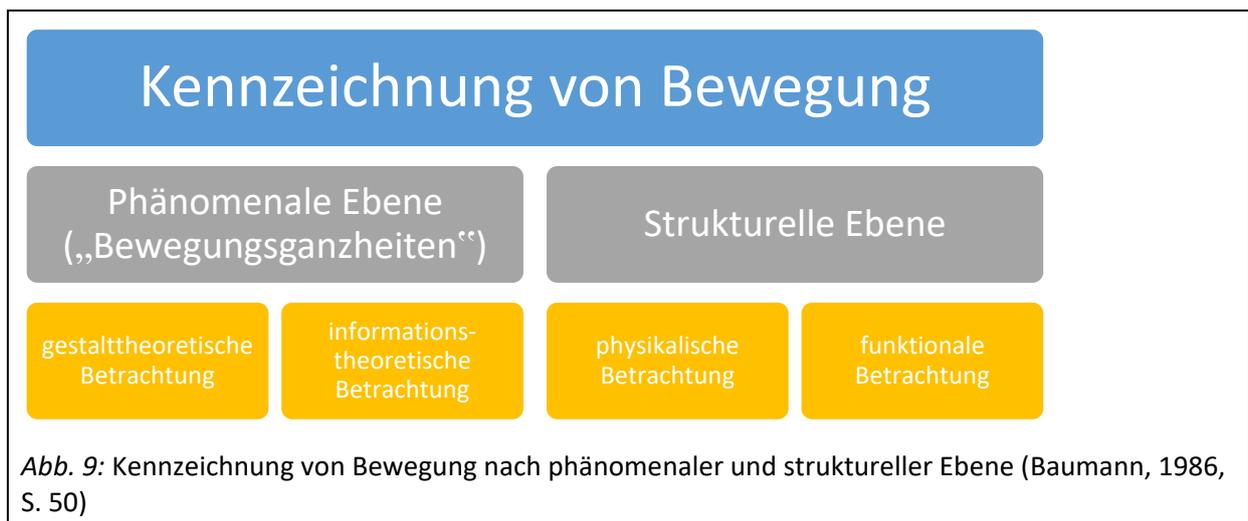


Abb. 9: Kennzeichnung von Bewegung nach phänomenaler und struktureller Ebene (Baumann, 1986, S. 50)

⁵ „Am Beispiel des Kurzschwungs [im alpinen Skilauf, Anm. d. Verf.] sollte die Beschreibung von Teilkörperbewegungen (z. B. der Beine, des Rumpfes) isoliert und im Zusammenhang erfolgen sowie die Bewegung der ‚Figur‘ im ganzen charakterisiert werden durch Attribute wie z. B. ‚rhythmisch‘“ (Baumann (1986, S.53)).

⁶ Bewegungsmuster (sog. patterns) können bestimmte isolierbare Aktionen aus dem Bewegungsablauf sein (wie z. B. die Schulter-, Arm- und Handgelenksstreckung, die als Wurfbewegung zusammengefasst werden) oder auch den gesamten Bewegungsablauf betreffen, wenn z. B. bestimmte Bewegungsausschnitte als flüchtige Körperpositionen im Bewegungsverlauf identifiziert werden (Baumann (1986, S.55f.)).

Entscheidend dabei ist, ob der Beurteilungsgegenstand die Gesamtbewegung ist oder nur einzelne Details des Bewegungsablaufs umfasst (Neumaier, 1988, S. 210f.). Kaminski (1975, S. 56f.) hebt zwar hervor, dass Bewegungsfehler „nur im Kontext eines sie ‚übergreifenden‘ Geschehens zu bemerken und sinnvoll zu interpretieren“ sind, der gesamte Bewegungsablauf aber in **Beobachtungseinheiten** strukturiert werden muss, um ihn „überschaubar und als Kommunikationsgegenstand greifbar zu machen“. Wie groß diese Beobachtungseinheiten letztlich sein können oder sollen, hängt in nicht unerheblichem Maße vom Beurteilungszweck ab. Relativ große Beobachtungseinheiten machen den Beurteilungsgegenstand zwar gut überschaubar, sind jedoch beobachtungsmethodisch problematisch, weil sie für Interpretationen und subjektive Einschätzungen des Beurteilenden großen Raum lassen. Je kleiner die Beobachtungseinheiten sind, desto interpretationsfreier und faktenorientierter sind sie zwar, aber desto schwieriger wahrnehmbar werden sie auch für den Beobachter (Kaminski, 1975, S. 57f.). Es geht also prinzipiell nicht darum, „möglichst kleine Teile aus dem Gesamtablauf der Bewegung zu isolieren, sondern für die Aufgabenstellung **sinnvolle** Teile oder Einzelmerkmale zu finden, die eine **objektive** Bewertung zulassen“ (Neumaier, 1988, S. 211). Dies darf aber sinnesphysiologisch wie psychologisch zu keiner Überforderung der Wahrnehmungsfähigkeit des Beurteilers führen, was die Wahrnehmbarkeit dieser Merkmale als Einzelelemente und als Kombination (z. B. die Geschwindigkeit der Bewegung, die notwendige retinale Abbildungsgröße für eine Detailwahrnehmung, die Anzahl der zu berücksichtigenden Details und ihre zeitliche Nähe zueinander sowie ihre räumliche Entfernung voneinander) anbelangt (Neumaier, 1988, S. 211f.).

Hartmann (1999, S. 109–111) spricht von sog. **figuralen Merkmalen**, welche einen vornehmlich optischen Eindruck des äußeren Verlaufs einer Bewegung liefern und somit auf die räumlich-zeitliche Struktur und die Harmonie einer Bewegung abzielen. Dabei ist besonders die Klassifikation sportlicher Bewegungen nach ihrer Phasenstruktur (nach Meinel, 1960) hervorzuheben, welche eine logische und zweckentsprechende Ordnung von Bewegungen ermöglicht, die sich je nach Bedarf aus sportartspezifischer Sicht weiter ausdifferenzieren und strukturieren lässt. Der Zusammenhang der einzelnen Phasen einer sportlichen Bewegung zueinander hat dabei hohe methodische Relevanz. Besteht zum Beispiel eine resultative Relation (als Ergebnisbeziehung) zwischen Vorbereitungs- und Hauptphase, so bedeutet dies, dass die Verbesserung der Vorbereitungsphase auch zu einer besseren Gestaltung der Hauptphase führt (wie sich beispielsweise durch die Erhöhung der Anlaufgeschwindigkeit eine größere Sprungweite beim Weitsprung ergeben müsste). Besteht hingegen eine finale Relation (als Zweckbeziehung), so wird während der Vorbereitungsphase bereits der Zweck der Hauptphase vorweggenommen und die Vorbereitungsphase entsprechend angepasst (wie dies beispielsweise beim Hochsprung der Fall ist, bei dem es nicht um eine maximale, sondern eine möglichst optimale Anlaufgeschwindigkeit geht, damit der Absprungvorgang noch zweckmäßig ausgeführt werden kann). Dies wäre beim Training der Teilschritte unbedingt zu beachten. Eine kausale Relation (als ursächliche Beziehung) liegt immer dann vor, wenn aus einer Phase zwangsläufig die nächste Phase folgt (wie dies beispielsweise bei Landevorgängen der Fall ist, welche automatisch auf den Absprungvorgang folgt). Somit muss manchmal auch rückwärts gedacht werden, wenn beispielsweise die Hauptphase erst erlernt werden kann, wenn die Endphase (sozusagen als Ausklang der Hauptphase) sicher beherrscht wird (wie dies beispielsweise bei der Landung im Skisprung der Fall ist).

Gerade die **dynamischen Merkmale** einer Bewegung sind wichtig, um ihren dynamisch-zeitlichen Verlauf durch gedankliches Mitvollziehen nachzuempfinden. Typische Merkmale hierfür sind Bewegungsrhythmus, -fluss, -elastizität und -übertragung (bzw. Bewegungskopplung). Besonders der Bewegungsrhythmus kann als strukturergänzendes Merkmal gut beobachtet werden über die Verteilung der

Kraftimpulse im räumlich-zeitlichen Verlauf und hat auch in der sporttechnischen Ausbildung methodische Relevanz, z. B. über die Rhythmusübertragung durch Mitvollziehen (Carpenter-Effekt), die Ausnutzung der Funktion des Gruppenrhythmus, die Erfassung und Wiedergabe des Bewegungsrhythmus sowie die rhythmische Lehr- bzw. Bewegungsweise (Hartmann, 1999, S. 111).

Wissenschaftlich betrachtet hat die morphologische Betrachtungsweise von Bewegungen mit der Kritik zu kämpfen, dass sie unexakt, nur qualitativ, ja sogar unwissenschaftlich sei, weil eben die Wahrnehmung im Vergleich zur messbaren Bewegungserfassung wie der Biomechanik subjektiv geprägt ist. Jedoch besteht der Vorteil solch phänomenologischer Analysen darin, „dass der Mensch Wahrnehmungsqualitäten aus den wechselseitigen Bezügen von Verhalten und Situation, von Wahrnehmung, Kognition, Erlebnis und Bewegung produziert, die wiederum dem Messsystem verborgen bleiben“ (Rostock & Kirchner, 1999, S. 117f.). Phänomene wie Bewegungsfluss, Bewegungskopplung, Bewegungsausdruck oder Bewegungsharmonie können zwar in Einzelparametern gefasst werden, aber nur in ihrer Ganzheitlichkeit beschrieben und erlebt werden. Daher scheint die morphologische Betrachtungsweise pädagogisch-didaktisch gesehen einen höheren Nutzen zu haben als eine vom Lernenden nur zum Teil selbst wahrnehmbare und erlebbare biomechanische Analyse (Rostock & Kirchner, 1999, S. 118f.). Daher wird in dieser Studie die morphologische Betrachtungsweise zur Messung der Beurteilungsleistung herangezogen und überlegt, welche Voraussetzungen bezüglich der zu beobachtenden Bewegungsmerkmale erfüllt sein müssen, um dieser Art der Bewegungsanalyse auch im Messverfahren gerecht zu werden und den Beurteiler sinnesphysiologisch nicht zu überfordern.

2.1.3.3.2 Limitierende Faktoren diagnostischer Prozesse im Sport

Der Hauptgegenstand des Sportunterrichts, das Sich-Bewegen der Schülerinnen und Schüler, liegt nicht wie in anderen Fächern zumeist in schriftlicher Form auf dem Schreibtisch, sondern muss an Ort und Stelle innerhalb von Sekundenbruchteilen beobachtet und beurteilt werden. Wenn eine Beurteilung der Bewegungsqualität ohne zusätzlichen technischen Aufwand vorgenommen werden soll, muss man mit der Nichtwiederholbarkeit der beobachteten Bewegung zurechtkommen (Meinel et al., 2015, S. 381). Die Herausforderung besteht also darin, den sportlichen Leistungs- und Könnensstand der Schülerinnen und Schüler überhaupt feststellen zu können. Die sinnesphysiologischen Voraussetzungen und Beschränkungen für diese komplexe Wahrnehmungsaufgabe werden daher in Kapitel 2.2 beschrieben, um zu klären, welche Informationen ein Beobachter überhaupt realistisch aufnehmen kann. Für objektiv messbare Leistungen, die mit Stoppuhr oder Maßband erhoben werden können, ist dies noch einfach realisierbar. Schwieriger wird es schon, wenn die Leistung eines Schülers in Bezug auf die technische Ausführung einer Bewegung oder die Bewegungskoordination generell festgestellt werden soll. Hierzu muss die Sportlehrkraft in der Lage sein, diese Bewegung in Echtzeit kriteriumsbezogen zu beobachten und zu beurteilen, um dem Lernenden daraufhin Rückmeldung geben zu können, welche Kriterien gelungen sind oder noch verbessert werden können. Die hohe Bewegungsgeschwindigkeit oder Beschleunigung von Sportgeräten, die Gleichzeitigkeit verschiedener Bewegungsabläufe sowie die eingeschränkte Sichtbarkeit auf relevante Bewegungsteile stellen oftmals limitierende Faktoren bei der Beobachtung von sich bewegenden Personen und Sportgeräten dar (Rockmann & Bömermann, 2006, S. 108f.). Ebenso weist der zu beobachtende Verhaltensstrom eine erhebliche Ereignisdichte auf, ist oft in feinsten Details regelungsrelevant und entzieht sich in vielen raum-zeitlichen Anteilen dem visuellen sowie auditiven Zugriff des Beobachters (Kaminski, 1975, S. 49f.). Und nicht zuletzt läuft der

Sportunterricht nicht so geregelt ab wie in einem Klassenraum, sondern bringt vielfach unvorhersehbare Bewegungen und Standortveränderungen der Beobachtungsobjekte am Ort des Sportgeschehens (Turnhalle, Sportplatz o.ä.) mit sich (Hecker, 1975).

Meinel et al. (2015, S. 381) schlagen daher vor, dass durch eine entsprechende Organisation der Beobachtung derartige Einflüsse auf die Beobachtungs- und Beurteilungsqualität möglichst gering gehalten werden sollten, um Fehlrteile zu vermeiden. Maßnahmen hierfür wären ein Ortswechsel des Beobachters bei wiederholter Demonstration der Bewegung oder die Konzentration auf Bewegungsdetails, die für das Gelingen vorrangige Bedeutung haben. Vor allem aber durch den Einsatz von Aufzeichnungstechnik wie Videoaufnahmen lasse sich die Wiederholbarkeit der Beobachtung – gegebenenfalls auch in Zeitdehnung – sicherstellen und somit auch das Beobachtungs- und Beurteilungsergebnis präzisieren.

Erst auf Basis dieser eben beschriebenen, komplexen Leistungsfeststellung ist es möglich, ein diagnostisches Urteil darüber abzugeben, ob und mit welcher Bewegungsqualität die Lernenden bestimmte Aufgaben und Übungen auf Grundlage vorhandener Bewegungsfertigkeiten bewältigen können. Daher stellt die spontane Beobachtungs- und Beurteilungskompetenz der Sportlehrkraft (im Sinne einer morphologischen Bewegungsanalyse, siehe Kapitel 2.2 und 0) ein wesentliches Kriterium der Diagnosekompetenz im Sport dar, um Schülerleistungen im Bereich des Bewegungslernens überhaupt adäquat erfassen zu können.

2.1.3.3.3 Einflussfaktoren auf die Bewegungsbeurteilung

Die qualitative Methode der Bewegungsbeobachtung gilt im Sport als ein „anerkanntes und aussagenrelevantes **Diagnose-** (Kontroll- und Untersuchungs-)verfahren [...], kurzfristig und ohne größeren (mess-)technischen Aufwand lern- und trainingsrelevante Informationen zu erhalten und darauf das weitere methodische Vorgehen abzustimmen“ (Meinel et al., 2015, S. 376). Es handelt sich dabei um eine absichtlich initiierte Fremdbeobachtung von sportlichen Bewegungsabläufen, die bewusst geschieht und insbesondere auf die visuelle Wahrnehmung von bestimmten Bewegungsmerkmalen und deren Ausprägungsniveau gerichtet ist (ebd., S. 380). Die beobachteten Merkmale sind dabei immer von den Erfahrungen des Beobachters abhängig, weshalb man von der Subjektivität der Wahrnehmung spricht (Rockmann & Bömermann, 2006, S. 109). Durch einen Abgleich dieser subjektiv gewonnenen Informationen mit dem eigenen Wissen und den eigenen Erfahrungen lassen sich Einschätzungen treffen, wie zielgerichtet und zweckmäßig eine Bewegungsausführung war oder auf welcher Qualitätsstufe eine sporttechnische Fertigkeit ausgeführt wurde. Somit stellt die Beobachtungskompetenz von Lehrenden im Sport eine wesentliche Voraussetzung dar für die Erfassung und Beurteilung von Bewegungsqualitäten und bleibt trotz der immer diffiziler werdenden technischen Messinstrumente im Sport eine der Hauptmethoden in der täglichen diagnostischen Arbeit von Sportlehrern, Trainern und Therapeuten (Meinel et al., 2015, S. 376f.).

Aus diesem Grund hat schon Meinel (1960, S. 138) Anfang der 1960er Jahre darauf hingewiesen, dass der Bereich „Bewegungsanalyse – Bewegungsbeurteilung – Bewegungsanweisung [...] das Kernstück der motorisch-methodischen Denkarbeit des Sportpädagogen“ darstellt. Dabei gehe es für Sportlehrkräfte weniger um die Fähigkeit, ideale Bewegungsformen darzustellen, sondern wesentliche Bewegungsmerkmale einzuschätzen, die auf einen gut koordinierten Bewegungsablauf hindeuten. Dabei müssen zum einen über das Analysatorensystem hoch sensible Wahrnehmungs-, Empfindungs- und

Beobachtungsprozesse (siehe die oberen drei Aspekte in Abb. 10) ablaufen können, zum anderen eine fein differenzierte, mit eigener Bewegungserfahrung und entsprechendem Fachwissen angereicherte Bewegungsvorstellung (siehe die unteren drei Aspekte in Abb. 10) vorliegen, um durch einen Sollwert-Istwert-Vergleich eine objektive Bewegungsbeurteilung zustande zu bringen (Hartmann, 1999, S. 107). **Bewegungsbeurteilung** ist also dadurch charakterisiert, dass das Beobachtungsergebnis eine Bewertung der beobachteten Bewegung zum Ziel hat (z. B. in Form einer Bewegungskorrektur) (Neumaier & Jendrusch, 1999, S. 128).



Abb. 10: Bewegungsbeurteilung beeinflussende Aspekte (nach Hartmann, 1999, S. 107)

Grundsätzlich spielt auch das **Bewegungsgedächtnis** des Beurteilenden eine wichtige Rolle bei der Bewegungsbeobachtung, weil sowohl die bereits erwähnten subjektiven Bewegungserfahrungen als auch die idealisierten Bewegungsvorstellungen (im Sinne eines Bewegungsleitbildes oder bestimmter Techniknotenpunkten) abgespeichert und im Rahmen eines Sollwert-Istwert-Vergleichs mit der beobachteten Bewegung abgeglichen werden können. Daraus resultiert dann entweder eine Übereinstimmung, welche auf einen erfolgreichen Lernprozess und die Bestätigung des bisherigen Vorgehens (im Sinne von ‚weiter so‘) hindeutet, oder eine Diskrepanz, die Korrekturen notwendig macht oder die angewandte Methodik hinterfragen lässt (Meinel et al., 2015, S. 381). Auch Umfang und zeitliche Dauer der Speicherung von Informationen sind dabei limitierende Faktoren der Bewegungsbeobachtung. Um die Fähigkeit des Arbeitsgedächtnisses im Zusammenhang mit Bewegungsvollzügen als möglichen Einflussfaktor auf die Beurteilungsleistung zu erfassen und messbar zu machen, kommt in dieser Studie der Test zur Kontrolle der Bewegungsvorstellungsfähigkeit (TKBV) von Schott (2013) zum Einsatz (siehe Kapitel 4.3).

Generell kann festgehalten werden, dass die Bewegungsbeobachtung bei geschulten und erfahrenen Observanden „prozessorientiert, d. h. vorausschauend (antizipierend), bestimmte Bewegungsqualitäten und -situationen erwartend“ (Meinel et al., 2015, S. 381) erfolgt, weshalb Hotz (1997, S. 201) von einer „antizipatorischen Wahrnehmungskontrolle“ spricht, die den Beobachter offen für Erwartetes und zugleich Unerwartetes hält. Diese Kompetenz ist sowohl bei der Leistungsbewertung sportlicher Bewegungen als auch bei der Initiierung motorischer Lernprozesse aller sporttechnischen Fertigkeiten vonnöten.

2.1.3.3.4 Hierarchisches Ebenenmodell der Bewegungsanalyse

Das hierarchisch geordnete Ebenenmodell dient als Grundlage für weitere Überlegungen, wie Bewegungen grundsätzlich analysiert werden können, um daraus ein Untersuchungsinstrument für die Messung der Bewegungsbeobachtung und -beurteilung abzuleiten. Hartmann (1999, S. 112f.) unterscheidet dabei allgemeine und spezifische Struktur- und Güteaspekte (siehe Abb. 11), die dazu beitragen sollen, dass die Analyse und Beurteilung von Bewegungen sowie entsprechende Bewegungsanweisungen zweckentsprechend, verständlich und in einem gewissen Maße auch objektiv erfolgen können. Diese Ordnung kann beispielsweise auch hilfreich sein, wenn es darum geht, die fachdidaktische Diagnosekompetenz mithilfe geeigneter Trainingsmaßnahmen verbessern zu wollen.

Die ersten zwei Ebenen behandeln allgemeine und spezielle Strukturaspekte einer Bewegung, wobei nur wesentliche und die Spezifik der Bewegung charakterisierende Merkmale in die Bewegungsbeschreibung eingehen sollen. Auf Ebene 1 fließen dabei in eine funktionsorientierte Phasenstruktur sowohl kinematische (im Sinne einer räumlich-zeitlichen Struktur der Teilbewegungen) als auch dynamische Aspekte (im Sinne einer Struktur der Teilimpulse und Muskeldynamik) ein. Während sich die räumliche Dimension (vornehmlich als Lage- und Raumveränderungen eines Körpers oder seiner Teile im Verhältnis zu einem definierten Bezugssystem) über die Phasenstruktur gut beobachten und abschätzen lässt, sind die für die Bewegungsausführung benötigte Zeit sowie die dafür aufgewendete Kraft bzw. Energie kaum visuell beziehungsweise nur indirekt zugänglich. Dynamische Aspekte können allenfalls an den erzielten Folgen der Krafteinwirkungen, über den Bewegungsumfang oder die Bewegungsstärke sowie indirekt über das innere Mitvollziehen der Bewegung abgeschätzt werden (Meinel et al., 2015, S. 378), während in zeitlicher Hinsicht am besten Geschwindigkeitsänderungen wahrgenommen werden können (Neumaier, 1988, S. 325).

Ebene 2 ist durch die Beschreibung des Bewegungsrhythmus sowie unterschiedlicher Arten der Bewegungskopplung (bzw. Bewegungsübertragung) gekennzeichnet. Bewegungsrhythmus ist gekennzeichnet von einer ausgewogenen Spannungsregulation der an einer Bewegung beteiligten Muskeln, wohingegen sich Koordinationsschwächen in diesem Bereich in einem schwerfälligen Impulswechsel sowie in einer unausgewogenen Vorspannung der antagonistischen Muskelgruppen zeigen (Hartmann, 1999, S. 114). Für die in dieser Studie zu beurteilende Wurfbewegung stellt der richtige Bewegungsrhythmus vor allem im Bereich der letzten drei Schritte einen leistungsbestimmenden Parameter dar, der einem Beobachter bekannt sein soll, um die qualitative Ausführung des Anlaufs zum Wurf beurteilen zu können.

Ebene 3 zielt auf die Wahrnehmung der allgemeinen qualitativen Güteaspekte Bewegungsfluss, Bewegungspräzision und Bewegungskonstanz ab und stellt die Übergangphase zu einer spezifischeren Bewegungsanalyse in Ebene 4 darstellt. Wirken Bewegungen unflüssig, so ist dies auf ein unausgewogenes Zeitmaß zurückzuführen und äußert sich in unangemessen schnellen, abrupten Impulsen oder verlangsamt, zähflüssigen Bewegungen bei zum Teil verzögerter Reaktion. Bewegungspräzision meint die Genauigkeit der Bewegungsausführung und ist immer dann im Raummaß als unausgewogen wahrzunehmen, wenn gleichgewichtsunsichere, aber auch zitternde, ausfahrende, eckige oder abgehackte Bewegungen auftreten (Hartmann, 1999, S. 114). Göhner (1984, S. 96) weist dabei darauf hin, dass auf dem Lernstand von Anfängern (entspricht Ebenen 1 bis 3) nur einfach zu verstehende Bewegungs-, Vergleichs- und Verlaufsbeschreibungen gegeben werden sollten, die folgende Merkmale beinhalten:

- räumliche (Angabe von Bewegungsrichtungen, wie z. B. vor-, rück-, seitwärts).
- zeitliche (z. B. während, danach, gleichzeitig, kurz),
- dynamische Parameter (z. B. kräftig, stark, sprunghaft) sowie
- Aussagen zur Ausgangs- und Endstellung.

Darüber hinaus gehende, überflüssige Bewegungsinformationen würden den Lernanfänger in seiner Bewegungsvorstellung überfordern.

Wenn es nun darum geht, konkrete Bewegungsanweisungen oder Korrekturhinweise zu geben, dann reichen diese allgemeinen qualitativen Einschätzungen nicht mehr aus, sondern müssen durch spezifische Güteaspekte (Ebene 4) ergänzt werden, wie dies beispielsweise bei sportart- bzw. disziplinspezifischen Schwerpunkten und Techniknotenpunkten der Fall ist, um tiefer in das Beziehungsgefüge der Strukturelemente einer Bewegung eindringen zu können. Dies setzt voraus, dass der Beurteilende über grundlegende Kenntnisse zu biomechanischen Wirkmechanismen und Gesetzmäßigkeiten verfügt, um das Funktionsgefüge (im Sinne von funktionstragenden Aktionen nach Baumann (1986, S. 59f.) richtig erkennen und analysieren zu können (Meinel et al., 2015, S. 379f.).

Auf der letzten Ebene 5 wird der Blick auf die Innensicht des Sich-Bewegenden gerichtet, indem psychische Aspekte des Handelnden als Ausdrucksformen der Verhaltensregulation bewertet werden. Meinel (1960) versteht darunter die bewusste Steuerung und Regelung einer Bewegung im Sinne der erreichten Bewegungskoordination, während in aktuelleren Forschungen auch psychische Befindlichkeiten und Einstellungen beim und zum sportlichen Handeln⁷ untersucht werden. Rostock und Kirchner (1999, S. 123–126) sprechen von sog. Handlungsmerkmalen der Orientierungs-, Antriebs- und Ausführungsregulation, von denen auf Ursachen, Absichten und Bewertungen des Handelns geschlossen werden kann (vgl. auch Rostock, 2003, S. 282).

Aus diesem Modell lässt sich ableiten, dass für die Bewegungsbeurteilung sowohl allgemeine als auch spezielle Strukturaspekte bekannt sein müssen, um auch die qualitativen Güteaspekte einer Bewegung sinnvoll beurteilen zu können. Diese kognitiven Leistungsdispositionen werden als ein möglicher Einflussfaktor auf die Beurteilungsleistung im Rahmen eines Fachwissenstests (siehe Kapitel 4.2) überprüft. Während die allgemeinen Strukturaspekte des Schlagwurfs hauptsächlich im Teil A des Fachwissenstests über den räumlich-zeitlichen Bewegungsverlauf sowie den Teil D zur visuellen Bewegungsvorstellung (über eine Bildkartensortieraufgabe) abgebildet sind, werden die speziellen Strukturaspekte und die speziellen qualitativen Güteaspekte verstärkt im Teil B über das biomechanische Wissen abgeprüft.

⁷ „Derartige Ausdrucksformen können z. B. sein: eine *erhöhte Erwartungshaltung* in Vorbereitung der Handlungsausführung, ein *gesteigertes Selbstbewusstsein* vor, während und nach der Handlungsausführung, das zunehmend *verminderte Selbstgespräch* vor und während der Handlungsausführung, eine *„gelöste Handlungsfreude“* auf der Grundlage einer fortgeschrittenen Automation in der Handlungsausführung (Dominanz der sensorischen Regulationsebene) und eine *verbesserte Variationsfähigkeit* in der Handlungsausführung insbesondere unter situativ-ändernden Bedingungen“ (Hartmann (1999, S.111f.)).

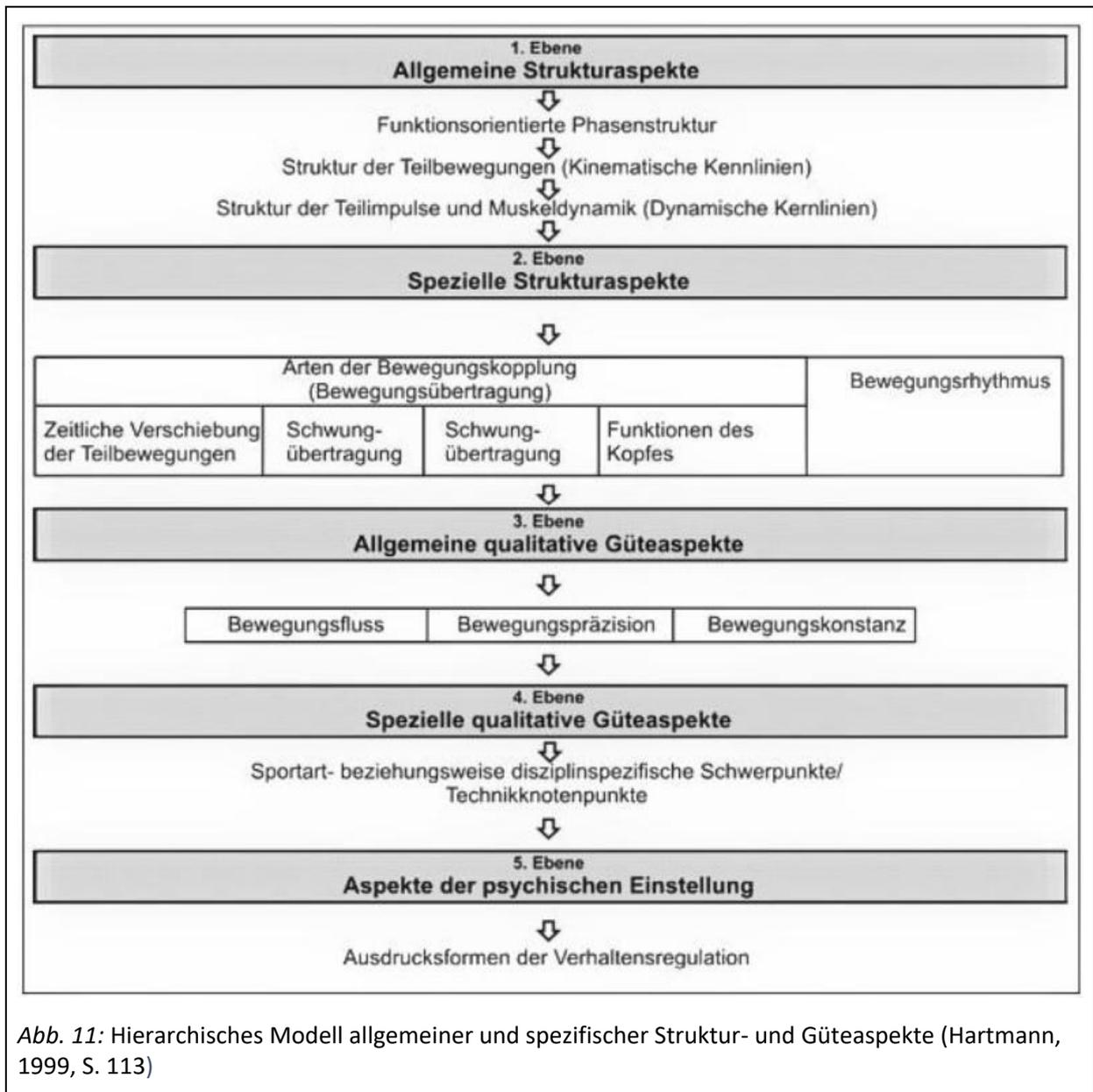


Abb. 11: Hierarchisches Modell allgemeiner und spezifischer Struktur- und Güteaspekte (Hartmann, 1999, S. 113)

2.1.3.3.5 Methodische Arbeit mit den Meinelschen Bewegungsmerkmalen

Die allgemeinen Meinelschen Bewegungsmerkmale⁸, welche in Hartmanns Ebenenmodell (siehe Abb. 11) in den Ebenen 1 bis 3 abgebildet werden, bilden zwar einen allgemeinen Bezugsrahmen, müssen aber um technikspezifische Beobachtungszugänge ergänzt werden, um eine sportliche Technik sinnvoll und umfassend beobachten zu können. Diese Erkenntnis ist eine wichtige Voraussetzung bei der Konstruktion des Beurteilungstests, welcher auf der Grundlage eines Beurteilungsrasters erfolgt (siehe Kapitel 4.5). Dabei dienen die Meinelschen Bewegungsmerkmale zwar als strukturgebende Beobachtungseinheiten, werden aber mit technikspezifischen Bewegungsdetails konkretisiert, um die verschiedenen Qualitätsstufen der einzelnen Bewegungsmerkmale beobachtbar zu machen. Dabei finden die

⁸ Gemeint sind nach Meinel et al. (2015, S.75–143) 1. Struktur sportlicher Bewegungen, 2. Bewegungsrhythmus, 3. Bewegungskopplung, 4. Bewegungsfluss, 5. Bewegungspräzision, 6. Bewegungskonstanz, 7. Bewegungsumfang, 8. Bewegungstempo und 9. Bewegungsstärke.

folgenden Erläuterungen von Rostock und Kirchner (1999, S. 120–123) unter anderem bei der Auswahl der zu beurteilenden Bewegungsmerkmale des Schlagwurfs Berücksichtigung:

1. Die Beobachtungszugänge über die Bewegungsmerkmale sollen die Spezifik der Bewegung widerspiegeln, wobei nicht alle, sondern nur die absolut charakteristischen Merkmale auszuwählen sind. Dabei ist zunächst von äußerlich sichtbaren Raum- und Zeitmerkmalen und ihren Veränderungen auszugehen (z. B. die kinematische Struktur einer Bewegung, welche die Bewegungsgeschwindigkeit, -umfang, -richtung, Zeitdauer, Timing und Körperlagen/-haltungen umfasst), wohingegen die Grunddimension Kraft (Bewegungsstärke) nur an den durch sie verursachten Wirkungen (z. B. als Bewegungsumfang anhand von Schrittlängen, Schwungweiten und typischen Ganzkörperstreckungen oder als Bewegungstempo anhand von Geschwindigkeiten) beobachtet werden kann. Danach ist die Qualität des Zusammenspiels von Einzelparametern (z. B. inwieweit dieses Zusammenspiel in Raum und Zeit fließend, rhythmisch, genau und gekoppelt erfolgt) zu prüfen.

2. Die Analyse mittels Bewegungsmerkmalen setzt eine ganzheitliche Bewegungsbetrachtung voraus, in welcher die Merkmale nicht isoliert und formal, sondern auch gekoppelt einzusetzen sind. Elementare Merkmale (z. B. Tempo, Stärke, Umfang) sollen zur Charakterisierung der komplexen Merkmale (z. B. Rhythmus, Kopplung) herangezogen werden

3. Das Ziel, welches durch die Bewegungsbeobachtung verfolgt werden soll, spielt ebenfalls eine Rolle. Grundsätzlich lässt sich festhalten, dass bei Bewegungsbewertung und -beurteilung häufig die ganzheitliche Betrachtung des Bewegungsablaufs im Mittelpunkt steht, während bei der Fehlerkorrektur der Blick stärker auf Details der Bewegungsausführung gerichtet wird. Soll also eine Bewegungsbewertung erfolgen (d. h. inwieweit und mit welcher Qualität das Bewegungsziel erreicht wurde), dann sind bei vollzugsorientierten Bewegungen vor allem der Bewegungsfluss, Bewegungsrhythmus und das Bewegungstempo miteinzubeziehen, während bei ergebnisorientierten Bewegungen das Erreichen des erwarteten Ziels (z. B. als Weite, Höhe, Punkte) einzuschätzen ist (z. B. als Ziel- und Treffgenauigkeit). Aufgabe des Beobachters ist es zu entscheiden, ob eine Bewegung flüssig, rhythmisch und zusammenpassend ist oder eher steif, eckig, unvollständig, unstimmig oder holprig. Gegensatzbeschreibungen wie „die Leichtigkeit und den Rhythmus der optisch wahrgenommenen Bewegung [mitspüren]“ im Vergleich zu „sich [...] von Ecken, Kanten, Stockungen, Entgleisungen und Abweichungen angestoßen“ (Kohl, 1990, S. 39) fühlen erleichtern dabei die Einschätzung. Wenn es um Fehlererkennung, Ursachenermittlung und Korrekturmaßnahmen geht, ist eine differenziertere und detailliertere Beobachtung erforderlich, indem in die Struktur einzudringen, das Zusammenspiel der Teilbewegungen zu erfassen sowie die Tatsache zu prüfen ist, inwieweit für jeweilige Ziele die erforderlichen Phasen schon ausgeprägt sind (Kohl, 1990, S. 38ff.).

4. Außerdem bilden die Bewegungskriterien eine Orientierungsgrundlage für die Abbildung der Innensicht, d. h. wie die Bewegung vom Handelnden selbst erfahren und erlebt wird. Dabei soll es zwischen der Außensicht des Bewerter und der Innensicht des Sportlers zu Übereinstimmungen kommen, weil der Lernende Korrekturanweisungen vor allem dann als wirksam einschätzt, wenn sie mit seinen eigenen Wahrnehmungen vergleichbar sind (Meinel, 1960, S. 361). Die Aufgabe des Lehrenden ist es also, bei der Korrektur von Fehlern zu prüfen, durch welche Wahrnehmungsmuster einzelne Bewegungsmerkmale (z. B. der Bewegungsrhythmus, -fluss, -umfang) erlebt werden können und wie diese Muster bei den Korrekturhinweisen wirkungsvoll verbalisiert werden können. Wenn Korrekturhinweise z. B. nur auf räumliche und zeitliche Phänomene des Vollzugs beschränkt blieben, die der Lehrende (abgesehen von filmischen Aufnahmen) nicht wahrgenommen hat, so würde dies die Effektivität der Fehler-

korrektur erheblich mindern. Auch der Zeitpunkt einer Bewegung, an dem kinästhetische Wahrnehmungsmuster (Ennenbach (1991, S. 186ff.) bezeichnet sie als Bewegungs- und Situationskerne) vom Subjekt besonders intensiv erlebt werden, können dabei ausgenutzt werden, beispielsweise bei:

- „Bewegungswechseln (Umkehr- und Wendepunkte, Richtungswechsel, Phasenübergänge)
- Geschwindigkeitsänderungen, Pausen und Abweichungen vom Timing oder bei
- Druckveränderungen, Spannungswechsel, Rhythmusveränderungen/-störungen“ (Rostock & Kirchner, 1999, S. 122f.)

2.2 Einflussfaktoren auf die Bewegungsbeobachtung

Grundsätzlich kann mit Erickson (2007, S. 19) für das Sich-Bewegen festgehalten werden: „There is little debate that vision is a critical factor in sports performance or that visual information is the dominant sensory system when performing practically any perceptual motor task such as those encountered in sports.“ Es erscheint einleuchtend, dass diese Betonung der Wahrnehmung auch auf die Personengruppe der Trainer oder Sportlehrer zutrifft, die hauptsächlich über den optischen Analysator Informationen über Bewegungen aufnehmen. Wenn man sich also mit der Beurteilung von Bewegungen beschäftigt, muss vorher geklärt werden, wie das Wahrnehmen von Bewegungen überhaupt funktioniert. Es stellt sich die Frage, wie der Mensch Informationen über Bewegungen sinnesphysiologisch aufnimmt und mit welchen Limitationen man zu rechnen hat, um bei der späteren Testkonstruktion die Beurteilungsaufgaben möglichst realistisch, aber nicht überfordernd zu gestalten. Daher wird in diesem Kapitel die Bewegungsbeobachtung im Hinblick auf die spätere Bedeutung für die Bewegungsbeurteilung näher beleuchtet, um daraus theoretische Erkenntnisse über mögliche Einflussgrößen auf die Beurteilungsfähigkeit zu gewinnen.

Die Bewegungsbeobachtung im Speziellen stellt sowohl für Sportlehrkräfte bei der Vermittlung motorischen Lernens als auch für Trainer im Rahmen des Techniktrainings und für Schieds-, Kampf- und Wertungsrichter zur Überwachung von Regelkonformität sowie zur Ermittlung der Bewertungsqualität eine wichtige Voraussetzung dar. Es könnte vermutet werden, dass es sich beim Bewegungssehen um eine allgemeine Fähigkeit handelt, die bei genannten Personenkreisen, die regelmäßig mit der Beobachtung sportlicher Bewegungen zu tun haben, besonders gut ausgeprägt ist. Jedoch kommen Biscan und Hofmann (1976) zu dem Ergebnis, dass erfahrene Sportlehrkräfte zwar bei der Beurteilung ihnen bekannter Bewegungen im Vergleich zu sportfremden Kollegen im Vorteil sind, hingegen bei völlig fremden Bewegungen (in diesem Falle eines frei erfundenen novel motor prototype) auch keine differenziertere Bewegungsbeobachtung zu leisten imstande sind. Daraus lässt sich vermuten, dass sich selbst durch berufsbezogene und gegebenenfalls jahrzehntelang ausgeübte Bewegungsbeobachtung keine generalisierbaren bewegungsanalytischen Fähigkeiten entwickeln, sondern es sich um eine hochspezialisierte visuell-dynamische Leistung handelt, die nicht einfach von einer Bewegung auf die andere übertragen werden kann. Jendrusch (2009, S. 130ff.) geht sogar so weit zu behaupten, dass Bewegungssehen nur sportart- bzw. bewegungsspezifisch trainiert werden kann, indem das Stimulusmaterial in der Übungsphase den Wahrnehmungsbedingungen in der Realsituation entspricht. Für die angestrebte Messung der Beurteilungsleistung werden daher in dieser Studie nicht verschiedene sportliche Bewegungen, sondern die leichtathletische Schlagwurfbewegung als konkreter Untersuchungsgegenstand ausgewählt. Die verwendeten Bewegungsbeispiele wurden außerdem im Rahmen eines schulischen Sportunterrichts gefilmt und stellen somit realistische Beurteilungsobjekte dar, mit denen die Probandinnen und Probanden als angehende Sportlehrkräfte im späteren Berufsalltag konfrontiert werden.

Dass es selbst unter erfahrenen Trainern bei der ad-hoc-Beurteilung eines vorgegebenen sportlichen Bewegungsablaufs zu deutlichen Divergenzen kommt, spricht umso mehr für die Tatsache, dass die Bewegungsbeobachtung eine grenzwertige Belastung oder sogar Überforderung für das visuelle System darstellt. Die beobachtungsbezogenen Anforderungen sind deshalb so hoch, weil es (zumindest ohne technische Hilfsmittel) keine Wiederholbarkeit des zu bewertenden Bewegungsablaufs gibt, Kriterien für Bewegungsqualität (z. B. Bewegungsfluss oder Rhythmus) schwer zu operationalisieren sind, eine hohe Ereignisdichte bei zum Teil sehr geringer Darbietungszeit vorliegt und bei hohen Objektgeschwindigkeiten oder Geschwindigkeitsänderungen mit einem zum Teil erheblichen Sehschärfeabfall zu rechnen ist. Wenn Bewegungen beobachtet werden sollten, die weniger als 200 Millisekunden dauern, dann kann es zu Wahrnehmungstäuschungen bezüglich der Bahnbewegung von Körperteilen kommen. Umso wichtiger ist es bei der Konstruktion eines geeigneten Testverfahrens für die Bewegungsbeurteilung, realistisch wahrnehmbare Beurteilungsaufgaben zu schaffen. Daher soll der Beurteilungstest dieser Studie verschiedene Schwierigkeitsstufen enthalten, um zu überprüfen, wie sich die zu untersuchenden Einflussfaktoren bei zunehmender Beurteilungskomplexität verhalten.

Morrison (2000, S. 23) vermutet beispielsweise, dass ein Beobachter besser in der Lage ist, eine Teilbewegung trotz ihrer Unbeobachtbarkeit richtig einzuschätzen, wenn er eine gute Kenntnis davon hat, welche Bewegungen biomechanisch bzw. anatomisch möglich sind. Anhand der Einflussgröße des biomechanischen Fachwissens kann später auf verschiedenen Komplexitätsstufen untersucht werden, ob sich dadurch möglicherweise bestimmte sinnesphysiologische Überforderungen kompensieren lassen. Dazu sollte man aber eine genaue Kenntnis über die visuell-dynamische Leistungsfähigkeit des optischen Systems sowie über die generellen Bedingungen der Bewegungsbeobachtung besitzen, weil die Bewegungsbeurteilung unter anderem von interpersonell variierenden Wahrnehmungsprozessen bestimmt wird.

Die Beobachtungskompetenz konstituiert sich dabei nach Tidow (1993, S. 17) aus drei Einflussfaktoren (siehe Abb. 1), deren inhaltliche Auseinandersetzung in den folgenden Unterkapiteln erfolgt:

1. **Bewegungs-Sehschärfe** mit ihren Subkategorien dynamische Sehschärfe und dynamische Ortungspräzision: gemeint ist die Fähigkeit des (ruhenden und sich bewegenden) Auges, sich bewegende Objekte (im zentralen und peripheren Sehbereich) möglichst präzise wahrnehmen zu können
2. **Blickmotorik**: ist geprägt von der Schnelligkeit und Präzision der Augenbewegungen
3. Verfügbarkeit und Realisierung bewegungsspezifischer Sehstrategien

Blickmotorik und Bewegungssehschärfe arbeiten dabei eng zusammen, weil es für die möglichst scharfe Wahrnehmung eines sich bewegenden Objektdetails unabdingbar ist, dass das Objekt auch bei hohen Geschwindigkeiten durch die Blickmotorik zielgenau verfolgt und möglichst präzise und konstant im Bereich des schärfsten Sehens gehalten wird. Daher werden beide Größen unter der Überschrift *Dynamische Sehschärfe* im Kapitel 2.2.1 gemeinsam dargestellt und ihre gegenseitigen Abhängigkeiten beleuchtet. Die sportsspezifische Sehstrategie wird in Kapitel 2.2.2 behandelt. Mit Helsen und Pauwels (1993, S. 385) bleibt aber auch zu bedenken, dass ‚Hardware-Komponenten‘ wie statische und dynamische Sehschärfe, peripherer Sehwinkelbereich oder die Zeitdauer visueller Korrekturen für die erfolgreiche Anwendung einer Sehstrategie eine geringere Rolle spielen als ‚Software-Komponenten‘ wie Informationsverarbeitungs- und Entscheidungsgeschwindigkeit oder Anzahl und Dauer der Fixationen.

2.2.1 Dynamische Sehschärfe (dynamic visual acuity)

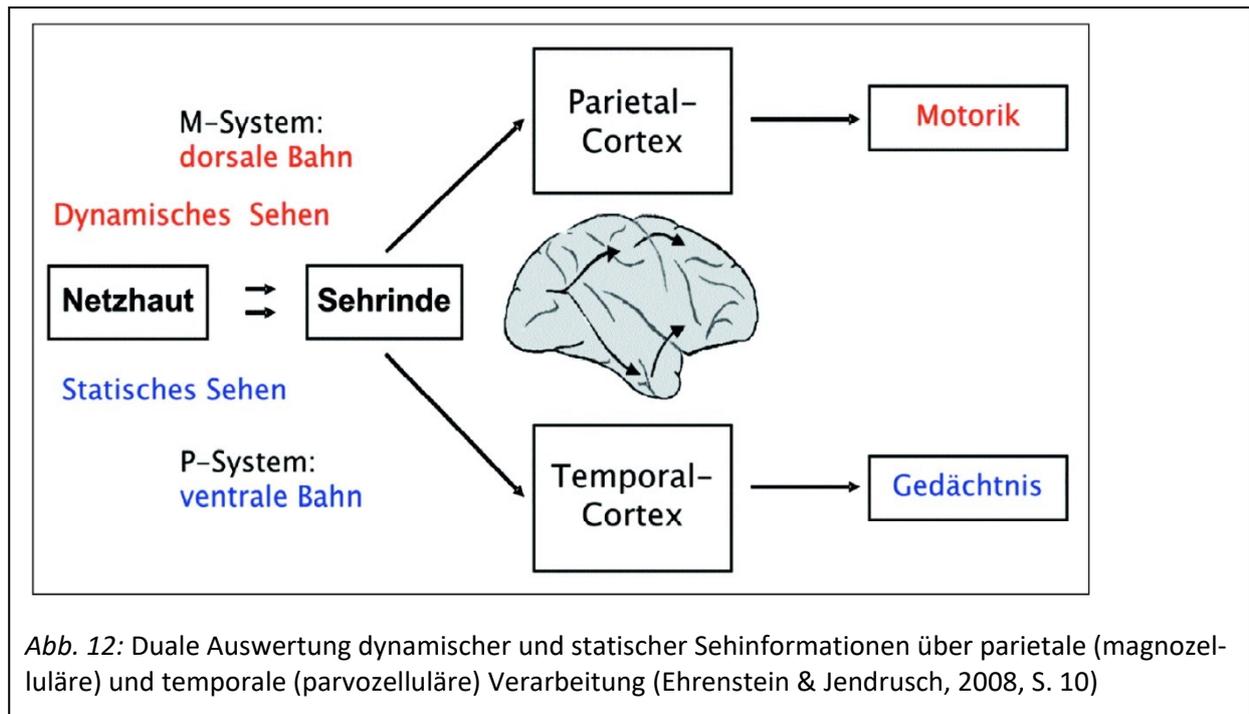
Generell ist das optische System aufgrund seiner großen Bedeutung für die Alltagsbewältigung des Menschen sehr leistungsfähig. Phylogenetisch betrachtet war es für den Menschen überlebenswichtig, Gefahren schnell zu erkennen und darauf adäquat zu reagieren. Vor allem das **periphere Sehen** spielt dabei eine entscheidende Rolle, indem es trotz eines schlechten räumlichen Auflösungsvermögens der Netzhautperipherie (Retina) sehr empfindsam gegenüber Bewegungen im Gesichtsfeldbereich⁹ ist und als Weitwinkel-Frühwarnsystem die Aufmerksamkeit durch ihre hohe zeitliche Auflösungsfähigkeit über das sog. transient system schnell auf solche Stellen im visuellen Raum lenkt, die potentiell neue Informationen liefern könnten, und gleichzeitig die Verarbeitungsprozesse des zentralen Sehens (im sog. sustained system) hemmt (Neumaier, 1979, S. 42–44). Die **dynamische Sehschärfe**¹⁰ (der Begriff dynamic visual acuity wurde erstmals von Ludvigh und Miller (1953) verwendet) beschreibt dabei die Fähigkeit, bewegte Objekte oder stationäre Objekte während Augenbewegungen (d. h. wenn eine relative Bewegung zwischen dem Beobachter und dem Beobachtungsobjekt vorliegt) möglichst scharf wahrnehmen zu können (Haase & Bossow, 2004, S. 79). Grundsätzlich herrscht zwar eine gewisse Abhängigkeit zwischen statischer Sehschärfe (d. h. dem möglichst scharfen Wahrnehmen eines unbewegten Objekts) und dynamischer, jedoch nimmt diese Korrelation mit zunehmender Bewegungsgeschwindigkeit des Beobachtungsobjektes ab (Erickson, 2007, S. 48).

2.2.1.1 Physiologische Grundlagen

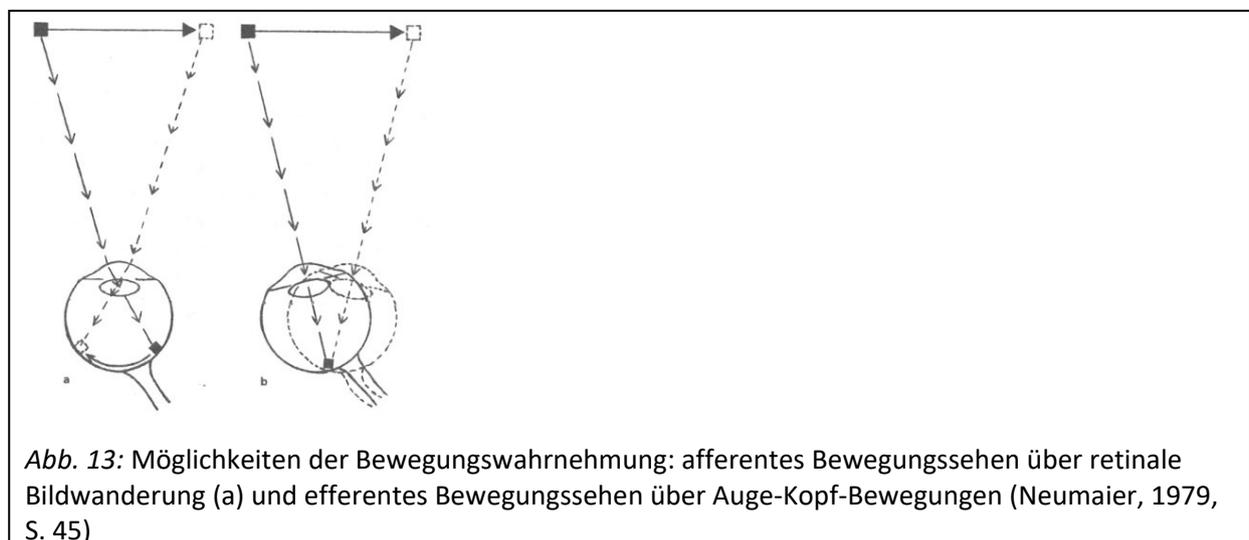
Es gilt im Folgenden die physiologischen Grundlagen zu verstehen, die beim statischen und dynamischen Sehen eine Rolle spielen. Neurowissenschaftliche Studien bestätigen, dass Sehen keine einheitliche Körperfunktion darstellt, sondern statisches und dynamisches Sehen durch zwei relativ unabhängige Teilsysteme verwirklicht wird: das kleinzellige oder parvozelluläre (P-)System dient vornehmlich dem Sehen von statischen Details, während das großzellige oder magnozellanuläre (M-)System für Raumorientierung und Bewegungssehen zuständig ist. Diese unterschiedlich spezialisierten Neuronensysteme sind bereits auf der Netzhaut unterscheidbar, wobei die M-Neuronen ein geringes räumliches, jedoch gutes zeitliches Auflösungsvermögen und die P-Neuronen umgekehrt eine hohe räumliche, aber nur geringe zeitliche Auflösung zeigen (siehe Abb. 12). Beide Systeme durchziehen die gesamte Sehbahn bis hin zur primären Sehrinde (Hinterhauptsbereich), wo sie sich dann verzweigen in eine temporale (Schläfenbereich) und eine parietale Verarbeitung (Scheitelbereich) und einen Übergang vom sensorischen auf das motorische visuelle System bewirken, indem sie beispielsweise Folgebewegungen in Gang bringen (Ehrenstein und Jendrusch (2008, S. 10); Haase und Bossow (2004, S. 67)).

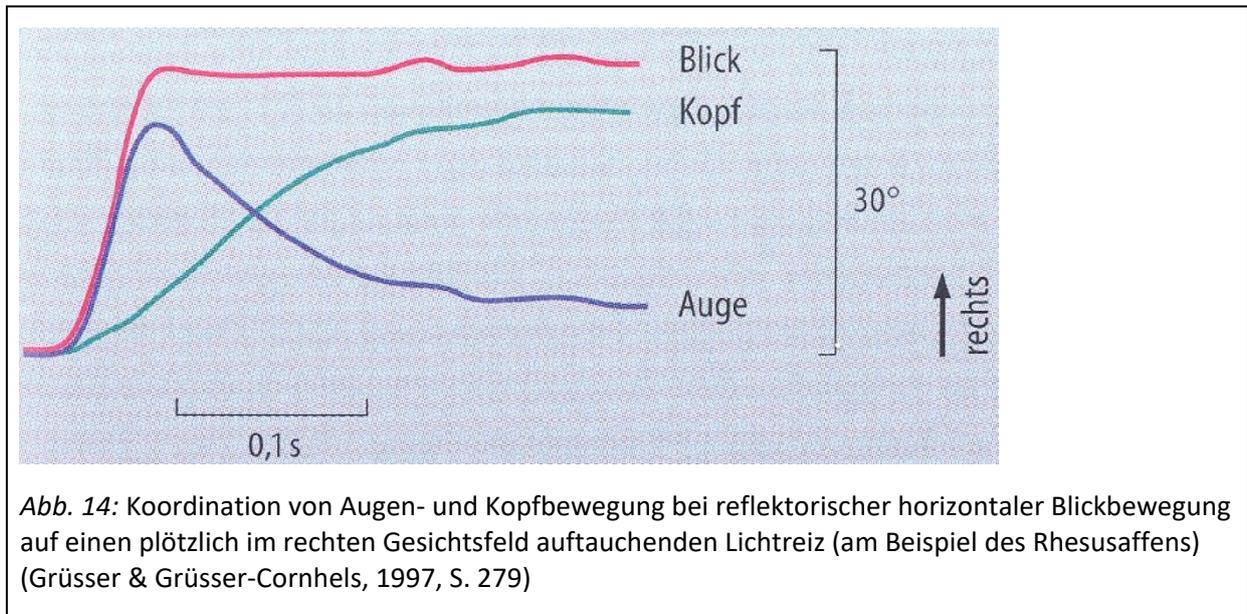
⁹ Das Gesichtsfeld definiert jenen Raumausschnitt, den das unbewegte Auge übersehen kann, und hat eine größere horizontale (ca. 180°) als vertikale (ca. 150°) Ausdehnung. Der Begriff Blickfeld grenzt sich dahingehend ab, dass sich der beobachtbare Raumausschnitt (immer noch bei ruhendem Kopf) mit Hilfe von Augenbewegungen erweitern lässt.

¹⁰ Haase und Bossow (2004, S.79) grenzen von der dynamischen Sehschärfe den Begriff **Bewegungssehschärfe** ab, welche die generelle Fähigkeit beschreibt, Bewegungen als solche wahrzunehmen. Tidow (1993) verwendet den Begriff Bewegungs-Sehschärfe (siehe Abb. 1) jedoch als Synonym für dynamische Sehschärfe.



Wird einem Objekt hohe Aufmerksamkeit geschenkt, so wird dieses mittels koordinierter Körper-, Kopf- und Augenbewegungen in den Bereich des schärfsten Sehens innerhalb der Fovea centralis (auch als Visus bezeichnet) gebracht und der Versuch unternommen, das Objekt durch Blicknachfolgebewegungen zu fixieren (**efferentes Bewegungssehen**, siehe Abb. 13, Bild b). Der Stimulus bleibt also durch das Auge-Kopf-Bewegungssystem auf derselben Stelle der Retina abgebildet, wobei die Augen zunächst schneller bewegt werden als der Kopf bzw. Körper. Indem der Kopf der Augenbewegung (diese wird über vestibuläre und mechanorezeptorische Signale der Halsbewegung sogar zurückbewegt) folgt, bleibt der Blick auch bei größeren Blickamplituden unbewegt auf das Objekt gerichtet (Höner, 2003, S. 101). Diese feinabgestimmte Kopplung zwischen Kopf- und Augenbewegung ist in Abb. 13 dargestellt. Auslösender Reiz für dieses efferente blickmotorische Einfangen und Verfolgen des sich bewegenden Objektes ist die sog. retinale Bildwanderung, bei der die Netzhaut durch die Bewegung des Objekts nacheinander an verschiedenen Stellen gereizt und somit eine Bewegung wahrgenommen wird, ohne dass das Auge oder der Kopf sich bewegt (**afferentes Bewegungssehen**, siehe Abb. 13, Bild a) (Mewes, Kellermann, Ehrenstein & Jendrusch, 2008, S. 141f.).





Fixation bedeutet demnach, dass das Auge ohne Rücksicht darauf, ob sich der Kopf oder der Körper (oder beide zusammen) bewegen (siehe Abb. 14), auf einen relativ eng umschriebenen Punkt im visuellen Reizfeld ausgerichtet ist und dort für eine gewisse Zeitspanne (ca. 0,3 s oder länger) haften bleibt.¹¹ Dabei liegt laut Neumaier (1979, S. 40–42; 46) der hundertprozentige Schärfebereich des zentralen Sehens nur innerhalb eines Winkelbereichs von 1°. Bei einer Objektabweichung von 3° zum Fixationspunkt reduziert sich die Sehschärfe bereits auf 50 Prozent, bei 6–7° Abweichung auf nur noch 25 Prozent (siehe Abb. 15). Angesichts dieser Tatsache, dass die Sehschärfe „aufgrund der abnehmenden Sensorendichte (Zapfen), der Signalkonvergenz (größere rezeptive Felder) sowie der sphärischen Aberration“ (Jendrusch, 2006, S. 57) zur Peripherie hin rapide abnimmt, wird verständlich, dass das Verfolgen und Fixieren bewegter Objekte, wie dies bei sportlichen Bewegungen zumeist der Fall ist, noch schwieriger ist als unter statischen Sehbedingungen und entscheidend von den blickmotorischen Fertigkeiten des Beobachters abhängt (siehe Kapitel 2.3.1). Eysel (2007, S. 387) bezeichnet dieses Phänomen, das nicht willentlich erzeugbar ist, als Augenfolgebewegung (Neumaier (1979, S. 38f.) verwendet dafür den Begriff **Blicknachfolgebewegung**) und meint damit kontinuierliche Gleitbewegungen des Auges, die dazu dienen, das Bild eines sich bewegenden Objektes auf der Retina möglichst stationär zu halten (siehe Abb. 13, Bild b). Sie benötigt eine gewisse Vorbereitungszeit des Auges (nach Erickson (2007, S. 57) mindestens 125 ms) und setzt erst ein, wenn die Winkelgeschwindigkeit eines Objektes die des Augendriftens während der Fixation erreicht.¹² Bis zu Winkelgeschwindigkeiten von 30–40°/s, bei geübten Sportlern sogar bis 100°/s (wenn die Bahnkurve des Objekts vorhersehbar ist), wird somit eine zufriedenstellende Wahrnehmung über glatte Folgebewegungen erreicht (Schnell (2010, S. 210); Erickson (2007, S. 58)), wohingegen bei höheren Geschwindigkeiten das Auge dann auf Sakkadenbewegungen (rapide Blicksprünge mit stark eingeschränkter visueller Aufnahmekapazität durch Suppressionsvorgänge) zurückgreift, die wiederum von kurzen, dazwischenliegenden Perioden von Folgebewegungen unterbrochen werden können. Bei solchen sog. **sakkadierten Folgebewegungen** kann die Folgebewegung bei großen Blickamplituden in der horizontalen Ebene Geschwindigkeiten bis zu 150°/s

¹¹ Physiologisch betrachtet ist die Fixation kein bewegungsloser Zustand der Augen, sondern diese überstreichen durch kleine Mikrobewegungen (small involuntary eye movements) in Form von Mikrosakkaden oder Gleitbewegungen (Drifts) fortwährend die Fixationsfläche (Haase & Bossow (2004, S.79)).

¹² Bis zu einer Winkelgeschwindigkeit von 2°/s kann ein Objekt durch Augenbewegungen ohne Herabsetzung der Sehschärfe verfolgt werden (Haase & Bossow (2004, S.79)).

erreichen (Haase & Bossow, 2004, S. 65f.). Bei Augenbewegungen von etwa $100^\circ/\text{s}$ verschlechtert sich das dynamische Sehvermögen schon deutlich und erreicht nur noch etwa zehn Prozent der statischen Sehschärfe (Marées & Mester, 1991, S. 148f.). So kann mit Erickson (2007, S. 48) festgehalten werden: „Visual acuity for a moving target is reduced compared with that of a stationary target and acuity becomes progressively more reduced with increasing velocity of the target.“

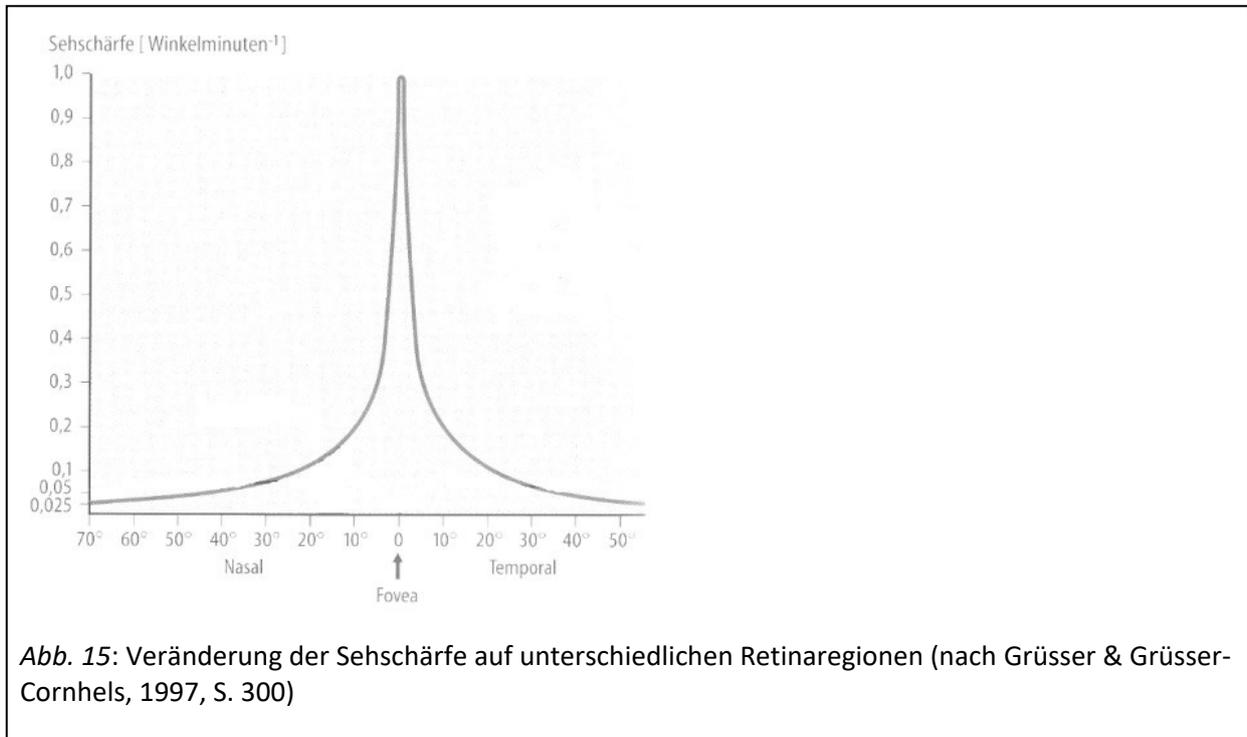


Abb. 15: Veränderung der Sehschärfe auf unterschiedlichen Retinaregionen (nach Grüsser & Grüsser-Cornhels, 1997, S. 300)

Sakkaden sind dabei ein wesentliches Merkmal der Beobachtung sportlicher Bewegungen und werden von Neumaier (1979, S. 36) als „rapide, ruckartige Augensprünge, durch die der Blick von einem Punkt zum anderen bewegt werden kann“ beschrieben. Dabei können maximale Winkelgeschwindigkeiten von etwa $600^\circ/\text{s}$ (Erickson (2007, S. 58) verweist sogar auf Messungen von bis zu $1000^\circ/\text{s}$) erreicht werden, wodurch allerdings in der Zeit, in welcher sich das Bild sakkadisch über die Oberfläche der Retina bewegt, durch Wahrnehmungssuppression keine visuellen Informationen aufgenommen werden können (Haase & Bossow (2004, S. 58); Weineck (2004, S. 103)).¹³ Eine Sakkade, welche weder in ihrer Richtung noch in ihrer Dauer willentlich beeinflusst werden kann, dauert in Abhängigkeit von ihrer Amplitude in der Regel nicht länger als 0,05–0,07 s bei einer Augenbewegung von ca. 20° . Weil bei Augenbewegungen über 15° entweder Kopfbewegungen einsetzen oder mehrere kleinere Sakkaden hintereinander erfolgen, beträgt die Sakkadendauer also zumeist nur zwischen 0,01–0,02 s. Die Latenzzeit vom Auftauchen eines visuellen Stimulus bis zum Beginn einer Sakkade beträgt zwischen 0,1–0,3 s (Erickson (2007, S. 57) gibt eine durchschnittliche Latenzzeit von 200 ms an, wobei diese Zeitspanne auch nicht durch sportliches Training verkürzt werden kann) und der zeitliche Abstand zwischen zwei Sakkaden erfordert 0,1–0,2 s, weshalb pro Sekunde nur 2–4 Sakkaden ausgelöst werden können (Neumaier, 1979, S. 36f.). Tidow (1996b, S. 265–267) bestätigt in einem sportspezifischen Sakkadenfrequenztest mit einer sehr großen Amplitude von 90° den Wert von 2,5 Sakkaden pro Sekunde. Da bei diesem Test aber Probanden mit Sakkaden-Geschwindigkeiten von über $600^\circ/\text{s}$ partiell geringere Frequenzen erzielten als solche mit nur $450^\circ/\text{s}$, darf angenommen werden, dass die Verweildauer

¹³ Die Wahrnehmungssuppression setzt sogar schon vor Beginn der Sakkade ein (Haase & Bossow (2004, S.79)).

bei der Fixation des Umkehrpunktes den eigentlich leistungslimitierenden Faktor darstellt. Je schneller die Informationsaufnahme und -verarbeitung im intersakkadischen Intervall erfolgt, desto früher wird der Reiz von der Fovea zu den höheren Verarbeitungszentren transferiert und eine neue Sakkade eingeleitet, weshalb Personen mit schnellerer Signal-Verarbeitung (und somit höherer Informationsverarbeitungskapazität) kürzer fixieren und somit eine höhere Maximalfrequenz erreichen. Kürzeste Verarbeitungszeit gepaart mit schnellster Blickmotorik stellt demnach eine zur Bewältigung von komplexen Bewegungsaufgaben ideale Konstellation dar. Erickson (2007, S. 58) spricht in diesem Zusammenhang von *inspection time*, welche ein Gradmesser für die Transfargeschwindigkeit der visuellen Informationen vom sensorischen Speicher in das Kurzzeitgedächtnis darstellt und somit für die Evaluation der Wahrnehmungsfähigkeit dienen kann. Dabei konnte Tidow (1996b, S. 272–277) nachweisen, dass sich durch ein Sakkadenfrequenztraining eine deutliche Verbesserung erzielen lässt, die nicht auf einer Erhöhung der Sakkadengeschwindigkeit beruht, sondern auf die Verkürzung der Verweildauer an den Umkehrpunkten um jeweils ungefähr 40 Prozent (von vorher 180 ms auf nachher 100 ms) zurückzuführen ist. Obwohl nicht direkt nachgewiesen werden konnte, dass sich durch die Steigerung der Blickfrequenz auch die visuelle Informationsverarbeitungszeit verkürzt, liegt die Vermutung nahe, dass die interpersonell variierende Fähigkeit zur visuellen Reizverarbeitung die Verweildauer am Fixationspunkt in hohem Maße determiniert und somit auch ‚mittrainiert‘ wird.

Bei hoher Winkelgeschwindigkeit eines Sehobjekts wird ein erfahrener Beobachter versuchen, den zeitweiligen Verlust des Blickkontakts durch Einsatz von Sakkaden wiederherzustellen und das Objekt in den Bereich des fovealen Sehens zu bringen, wodurch eine kontinuierliche Informationsaufnahme erschwert wird. Um Informationsverluste so gering wie möglich zu halten, spielt das sog. **antizipatorische Timing** eine wichtige Rolle, die Fähigkeit nämlich, „einen Blicksprung mit der (schnellen) Objektbewegung so zu koordinieren, daß [sic!] dieser zu einem Zeitpunkt gestartet wird, zu dem der angesteuerte Wahrnehmungsort noch ‚leer‘ ist“ (Tidow, 1993, S. 22). Das Auge muss also auf Basis der Bewegungsrichtung und Geschwindigkeit des Bewegungsgegenstands vorausberechnen, wo sich dieser in Relation zu dem für den Blicksprung benötigten Zeitbedarf befinden wird, um ihn zumindest kurzzeitig wieder foveal erfassen zu können, wobei keine Korrekturmöglichkeiten während der Sakkade möglich sind.

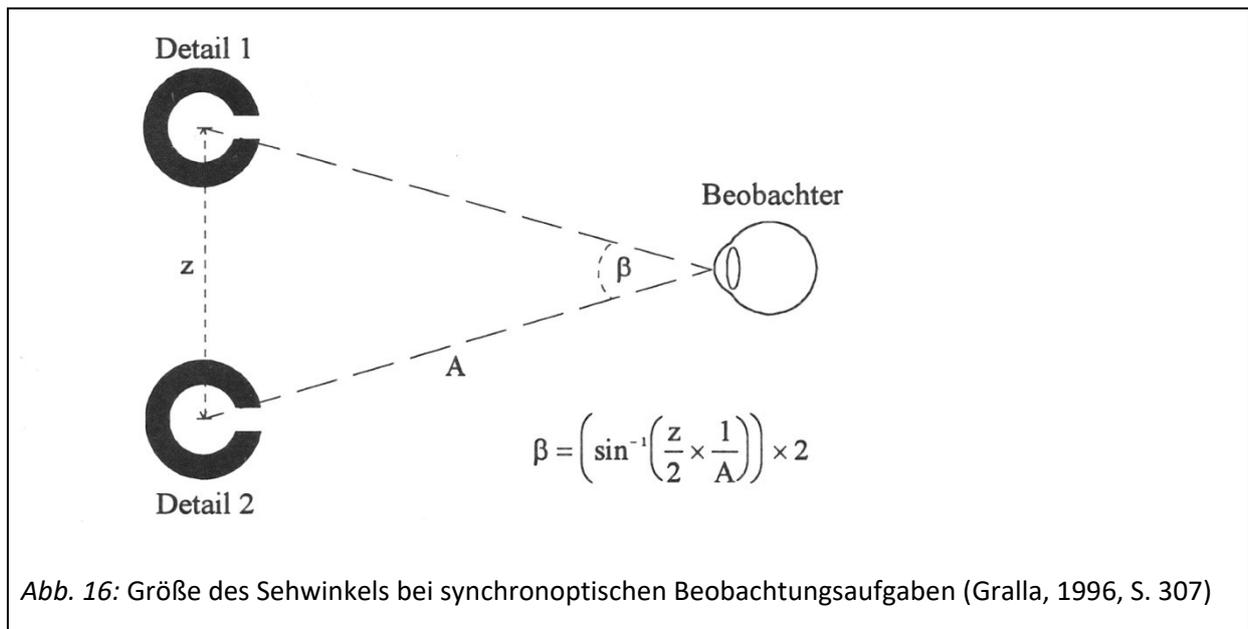
Die **Verarbeitung perzeptueller Informationen** selbst kann nach Morrison (2000, S. 23) in drei Subprozesse unterteilt werden: Organisation, Integration und Dekodierung. Wahrnehmungsfiler stellen dabei die erste Stufe des Verarbeitungsprozesses dar. Auf der Organisationsebene spielt die Fähigkeit, Teile einer Bewegung aus der Gesamtbewegung herausgelöst zu beobachten (*embedding*), eine entscheidende Rolle und beeinflusst die allgemeine Beobachtungs- und Analysefähigkeit, indem die Qualität der zu verarbeitenden Informationen für Analyseentscheidungen verbessert werden. Auf zweiter Ebene werden Informationen aus unterschiedlichen Sinneswahrnehmungen koordiniert und in ein gemeinsames Wahrnehmungsergebnis integriert. Schwierigkeiten in dieser Phase können Fehlurteile der beobachteten Bewegung verursachen. Auf der letzten Stufe der Dekodierung werden die aufgenommenen Reize mit Bedeutung angereichert, indem bereits bekannte Reizmuster aus dem Gedächtnis mit der eben aufgenommenen Wahrnehmungsinformation in Beziehung gesetzt werden (Biscan & Hofmann, 1976). Falls solche Vergleichsmomente aus eigenen Lehr- oder Bewegungserfahrungen heraus vorliegen, kann die Bewegungsanalyse auch nach allgemein-logischen Analysegesichtspunkten auf der Grundlage biomechanischer Prinzipien wie z. B. im B-BOAT-Konzept von (Abendroth-Smith & Kras, 1999) erfolgen. Daher können gerade biomechanische Informationen über eine Bewegungsausführung hilfreich sein, wenn die Eigenerfahrungen über die Bewegung nur gering sind.

2.2.1.2 Synchronoptisches Sehvermögen

Sollen bei komplexen Beurteilungsaufgaben auch noch zwei räumlich voneinander getrennte Aspekte gleichzeitig beobachtet und präzise erfasst werden (**synchronoptisches Sehvermögen** als Sonderform des peripheren Sehens; siehe Neumaier & Jendrusch (1999, 136f.)), dann ist der Ausprägungsgrad der Sehschärfereduktion außerdem noch abhängig von der Distanz (und damit der Sehwinkelgröße, siehe Abb. 16) zwischen den beiden kritischen Objekten, von der Größe sowie von der Winkelgeschwindigkeit der zu identifizierenden Objektdetails (Gralla, 1996, S. 307f.). Synchronoptisches Sehen stellt dabei eine wichtige Voraussetzung für die Beurteilung von Fremdbewegungen dar, wie dies an leichtathletischen Beispielen nachvollziehbar wird:

Wenn ein Trainer die Technik eines Kugelstoßers oder einer Hochspringerin beurteilen soll, müsste er eigentlich mehrere Körperpartien gleichzeitig beobachten. Weil dies aber nicht möglich ist, bedarf es optimaler Stellung und Sehstrategien, um nicht durch Sakkadensuppression wichtige Übungsphasen zu versäumen. (Schnell, 2010, S. 212)

Zwei Sehstrategien sind dabei grundsätzlich möglich: man kann das eine Objekt fixieren und die relative Unschärfe des anderen in Kauf nehmen, oder den Fixationspunkt zwischen die beiden Objekte setzen, wodurch keines von beiden im fovealen Bereich projiziert wird. Eine optische Sehschärfeminderung auf den rezeptiven Feldern für mindestens eines der Objekte ist in beiden Fällen zwangsläufig zu akzeptieren (Tidow, 1996b, S. 248f.). Eine weitere Unterscheidung der synchron-optischen Bewegungsbeurteilung ist bezüglich der Qualität des Synchronitätsurteils möglich. Bei zeitpunktbezogenen Urteilen muss entschieden werden, ob zu einem bestimmten Zeitpunkt zwei oder mehrere Bedingungen erfüllt sind oder nicht (z. B. ob beim Handballsprungwurf der Ball die Wurfhand verlassen hat, bevor der Fuß den Boden des Wurfkreises berührt), wohingegen zeitraumbezogene Urteile auf die verlaufsorientierte qualitative Beurteilung betreffender Übereinstimmungen mit einer biomechanisch-physiologischen Idealbewegung abzielen. Letztere synchronoptische Identifizierungsaufgaben sind inhaltlich anspruchsvoller und dienen dem Zwecke der Bewegungskorrektur, Optimierung von Fertigkeiten, Minimierung von Normabweichungen oder zur schulischen Notenbewertung und gehören zum Alltag jedes Sportlehrers und Trainers, wenn beispielsweise bei der turnerischen Kippbewegung am Reck zum Zeitpunkt des Kippstoßes zwei räumlich voneinander getrennte Haltungsausführungen, in diesem Falle das Strecken der Arme und das schnellkräftige Öffnen des Hüft-Beinwinkels, zu identifizieren und hinsichtlich ihrer Ausführungsqualität zeitgleich zu beurteilen sind (Gralla et al., 1998, 166f.). Die Trainierbarkeit des synchronoptischen Identifizierungsvermögens wird in einer Verbesserung der Aufmerksamkeitsleistung durch ungewohnte visuell periphere Belastungen vermutet, wobei eine Übertragung des Trainingserfolgs nur für Situationen mit vergleichbarem Anforderungscharakter gegeben ist, d. h. entsprechende synchronoptische sportartspezifische Beobachtungsanforderungen müssen - wenn möglich unter Verwendung entsprechender Korrekturinstanzen wie nachträglicher Videoüberprüfung - möglichst häufig bewältigt werden, um eine hohe Beurteilungssicherheit für die entsprechende Bewegung zu erlangen (Gralla et al., 1998, S. 169).



2.2.1.3 Messbarkeit und Trainierbarkeit

Schon 1953 wurde von Ludvigh und Miller (1953, zitiert nach Tidow, 1993, S. 31f.) ein standardisierter Test zur Ermittlung der dynamischen Sehschärfe vorgestellt, wobei dieser jedoch nicht besonders aussagekräftig für Anwendungssituationen im Sport ist, weil die Fähigkeit einer Person untersucht wird, bei gleichbleibender Winkelgeschwindigkeit ($100^\circ/\text{s}$) immer kleiner werdende Objektdetails (hier die Öffnung eines Landoltringes) zu identifizieren. Da es im Sport jedoch wesentlich häufiger vorkommt, dass Beobachter Sehobjekte mit unterschiedlichen und oftmals noch höheren Winkelgeschwindigkeiten wahrnehmen müssen, dafür aber weitgehend konstante Größe zeigen und nur selten höhere Visusanforderungen stellen, wurde der Testaufbau von Tidow (1993, S. 32–37) entsprechend für die Simulation sportlicher Beobachtungssituationen modifiziert. Es wurde die **maximale Ortungsgeschwindigkeit** (bei fixiertem Kopf) gemessen, indem stufenweise die Darbietungsgeschwindigkeit der Sehzeichen erhöht wurde. Da Objektbewegungen mit einer Winkelgeschwindigkeit über $100^\circ/\text{s}$ nicht mehr kontinuierlich über gleitende Augenbewegungen verfolgt, sondern kurzzeitige Fixationen des sich bewegenden Objekts über Blicksprünge (Sakkaden) hergestellt werden, wurde von den Versuchspersonen in diesem Test eine völlig neue Wahrnehmungstechnik gefordert, die schon eher einem sport-spezifischen Beobachtungsverhalten entspricht. Sportstudierende beispielsweise erreichten im Mittel $220^\circ/\text{s}$, wobei diese Leistung ohne Kopffixation noch deutlich gesteigert werden konnte. Interessant ist auch, dass bei Versuchspersonen, die visuell belastende Sportarten betrieben (z. B. Tennis- oder Badmintonspieler, Handballer, Skifahrer), eine höhere dynamische Sehschärfe nachgewiesen werden konnte (Ehrenstein & Jendrusch, 2008, S. 12f.). Bemerkenswert erscheint, dass die Leistungsfähigkeit auch von der Blickrichtung abhängig ist und nur ein geringer oder gar kein Transfer in der jeweiligen Gegenrichtung stattfindet. So fand Tidow (1996b, S. 262–265) heraus, dass unabhängig von Trainingseffekten die maximale Ortungsgeschwindigkeit von links nach rechts höher ist als von rechts nach links. Dieses Ergebnis hat auch Auswirkungen auf das Training des dynamischen Sehens, welches in beide Richtungen geübt werden muss, oder die richtige Wahl des Beobachterstandorts bei einer Bewegungsbeurteilung.

Die generelle Trainierbarkeit der blickmotorischen Leistungsfähigkeit, die in zahlreichen Studien (u.a. Tidow (1996b) und Banks, Moore, Liu & Wu (2004)) nachgewiesen werden konnte, wird „ursächlich weniger auf muskuläre als vielmehr auf zentral-koordinative Effekte zurückgeführt, d. h. auf verbesserte oder schneller abrufbare Motorikprogramme aus den Zentren für die horizontale Blickmotorik in der pontinen *Formatio reticularis*“ (Ehrenstein & Jendrusch, 2008, S. 13). Speziell für die Zielgruppe der Sportlehrer und Trainer ist das visuell-dynamische Leistungsvermögen eine entscheidende Einflussgröße, kommt es doch bei der Bewegungsbeobachtung darauf an, die entscheidenden Informationen innerhalb von Sekundenbruchteilen erkennen und verarbeiten zu können. Zugleich bleibt aber anzuerkennen, dass die dynamisch-visuelle Sehschärfe nicht sportadäquat messbar ist, weshalb dieser Einflussfaktor in dieser Studie unberücksichtigt bleiben muss. Jedoch können und sollen Trainingsinstrumente zur Verbesserung der Bewegungsbeurteilung Inhalte enthalten, die auf die Verbesserung der blickmotorischen Fähigkeit abzielen, um die Beobachtung sportlicher Bewegungen zielgerichteter gestalten zu können.

2.2.1.4 Schlussfolgerungen für die Messung der Beurteilungsleistung

Es bleibt festzuhalten, dass die Bewegungsbeobachtung sinnesphysiologisch eine sehr anspruchsvolle Aufgabe ist, da es sich bewegende Objekte sind, die in der *Fovea centralis* gehalten werden müssen, um eine ausreichende Informationsaufnahme sicher zu stellen. Wenn das Objekt nicht mehr mittels Blicknachfolgebewegungen kontinuierlich verfolgt und fixiert werden kann, so greift das Auge auf Sakkadenbewegungen zurück, bei denen es zu einer starken Wahrnehmungssuppression kommt. Bei den langsamer ausgeführten Bewegungsphasen wie dem Anlauf oder der Wurfarmrückführung erscheint die Blicknachfolgebewegung noch realistisch zu sein, aber bei der eigentlichen Hauptphase der Armbeschleunigung ab dem Einnehmen der Wurfauslage ist eine detaillierte Informationsaufnahme schwierig oder bisweilen sogar unmöglich. Dieser Tatsache muss dahingehend Rechnung getragen werden, dass die vorgegebenen Beobachtungsmerkmale nicht auf zu konkrete Beobachtungsdetails (wie beispielsweise genaue Winkelstellungen von Körpergelenken oder Bahnverläufe von Körperteilen) zielen, sondern in einem größeren Bewegungskontext wahrgenommen werden können. Dies spricht dafür, umfassendere Kriterien wie beispielsweise die Meinel'schen Bewegungsmerkmale für die Messung der Beurteilungsleistung zu verwenden.

Auch das synchronoptische Sehvermögen spielt eine Rolle bei der Entwicklung der Beurteilungskriterien. Da es sinnesphysiologisch nicht möglich ist, weit auseinanderliegende Körperpositionen wie beispielsweise Wurfarm und Fußaufsatz des Stemmbeins gleichzeitig zu fixieren, muss darauf geachtet werden, dass die konstruierten Beurteilungsaufgaben auch realistisch beobachtbar sind. Auch hier scheint der Vorzug von übergreifenden Bewegungsmerkmalen auf der Hand zu liegen, weil die Beobachtung nicht anhand weniger Bewegungsdetails erfolgt, sondern die einzelnen Bewegungsmerkmale sich über längere Phasen der Gesamtbewegung erstrecken. So kann der Bewegungsumfang der Ausholbewegung beispielsweise schon während der Aktion der Wurfarmrückführung und der Hüftrotation beobachtet und antizipiert werden und muss nicht erst mit Einnehmen der Wurfauslage beurteilt werden. Auf diese Weise erscheint es durchaus möglich, dass mehrere Bewegungsmerkmale im Laufe der Gesamtbewegung beurteilt werden können, auch wenn sich einige Teile der Bewegungsphasen überschneiden und synchron ablaufen. Es kann auch vermutet werden, dass die Beurteilungsleistung abnimmt, wenn die Komplexität der Beurteilungsaufgaben steigt, indem beispielsweise mehrere Bewegungsmerkmale in einem Beobachtungsdurchgang beurteilt werden sollen, was wiederum auf eine synchronoptische Überforderung hindeuten kann.

Um die Beurteilungssituation möglichst realistisch abzubilden, soll die Bewegungsbeobachtung im Beurteilungstest auf efferente Weise erfolgen, indem auch Bewegungen von Auge, Kopf oder Körper für die Verfolgung des zu beurteilenden Athleten notwendig sind. Würde man eine zu kleine Wiedergabefläche (z. B. einen Laptopmonitor) verwenden, so würde vorwiegend eine afferente Bewegungswahrnehmung erfolgen, weil das Objekt hauptsächlich über die retinale Bildwanderung verfolgt werden kann. Daher wird der Beurteilungstests ausschließlich auf großen Smart-TVs durchgeführt. Da sich die Messbarkeit des dynamischen Bewegungssehens insgesamt aber technisch sehr aufwändig gestaltet, muss auf eine Beurteilung dieses Einflussfaktors verzichtet werden.

2.2.2 Sportspezifische Sehstrategien

Grundsätzlich ist es bei der Bewegungsbeobachtung hilfreich, wenn der Beobachtende auf der Grundlage der verfügbaren Vorinformationen über eine Bewegung eine systematische Beobachtungsstrategie (systematic observational strategy) entwickelt, um die beurteilungsrelevanten Merkmale und Details einer Bewegung zuverlässig wahrnehmen und die Leistungsfähigkeit des optischen Systems (mit ihrer Begrenztheit, siehe Kapitel 2.2.1) optimal ausnutzen zu können. Dabei geht es zunächst noch gar nicht darum, die Beobachtungen zu bewerten, sondern die große Fülle an Bewegungsmerkmalen auf diejenigen critical features zu reduzieren, die zum einen eine hohe Wertigkeit für den Bewegungsvollzug haben, zum anderen aber auch ohne technische Hilfsmittel beobachtbar sind. Dabei kommt der Sehstrategie des Beobachters eine entscheidende Bedeutung zu, um das Auge zur richtigen Zeit auf den richtigen Ort zu lenken und die relevanten Bewegungsmerkmale aufnehmen und verarbeiten zu können:

Grundlage der Interpretation des Blickverhaltens ist zunächst die Erkenntnis, daß [sic!] eine ‚differenziert-identifizierende Informationsaufnahme‘ [...] nur im Bereich des scharfen Sehens, also im Bereich der Blickfixation erfolgt. Damit kann der **Fixationsort** mit dem Ort der Informationsaufnahme [...] gleichgesetzt werden, wenn die Aufgabenstellung das Erkennen von Details verlangt [...], z. B. die Identifikation lernrelevanter Elemente von visuellen Lehrmaterialien oder Bewegungsdemonstrationen im Rahmen der Schulung der Bewegungsvorstellung oder auch bei der Identifikation von Bewegungsfehlern zum Zwecke der Bewegungskorrektur. (Neumaier, 1988, S. 246)

Zugleich bleibt aber auch zu bedenken, dass die Wahrnehmung ein aktiver Prozess ist, in welchen der Beobachter sein Vorwissen und seine Erfahrungen ebenso wie seine individuelle Beobachtungsstruktur miteinbringt, weshalb das gleiche Objekt von verschiedenen Beobachtern auch unterschiedlich wahrgenommen werden kann (Knudson, 2013, S. 98f.).

Die wissenschaftlichen Erkenntnisse über die sportspezifischen Sehstrategien können dann vor allem für die Konzeption des Beurteilungstests genutzt werden. Vor allem geht es um die Frage, unter welchen Rahmenbedingungen der Beurteilungstest stattfinden soll. Wie detailliert müssen die Beurteilungskriterien erklärt werden, damit die Testpersonen das gleiche inhaltliche Verständnis über die einzelnen Qualitätsstufen haben und somit auch einheitlich beurteilen können? Wieviele Beobachtungsdurchgänge sind nötig, um eine umfassende Beurteilung der Gesamtbewegung vorzunehmen? Wieviel Lenkung der Beobachtungsstrategie ist sinnvoll, um die Bewegungsbeobachtung der Testpersonen zu unterstützen?

Dafür wird zunächst eine Visual Task Analysis der Schlagwurfbewegung angefertigt, um sich bewusst zu machen, mit welchem optischen Anforderungsprofil die spätere Bewegungsbeurteilung erfolgt (siehe Kapitel 2.2.2.1). Diese allgemeinen Faktoren bieten schon eine erste Entscheidungshilfe, welche Sehstrategien sich bei der leichtathletischen Schlagwurfbewegung anbieten oder ausschließen. Etwas konkreter wird Knudson (2013), wenn er für eine systematische Beobachtungsstrategie fünf Schlüsselemente empfiehlt:

A good systematic observation strategy should include what critical features to focus attention on, how to control the situation, the vantage points of observation, the number of observations needed, and a decision on whether extended observation will be needed. (Knudson, 2013, S. 103)

Von diesen Überlegungen ausgehend werden im Folgenden die Wahl des Aufmerksamkeitsfokus (siehe Kapitel 2.2.2.2), die Gestaltung der Beobachtungssituation (siehe Kapitel 2.2.2.3) und die Beobachtungswiederholung (siehe Kapitel 2.2.2.4) näher betrachtet. Aus diesen Erkenntnissen werden dann Schlussfolgerungen für die Konstruktion von Beurteilungsaufgaben gezogen, die im Rahmen dieser Studie befolgt werden sollten (siehe Kapitel 4.5.3).

2.2.2.1 *Visual Task Analysis der Schlagwurfbewegung*

Bevor man sich über eine konkrete Beobachtungsstrategie Gedanken macht, kann über eine „**Visual Task Analysis**“ (Erickson, 2007, S. 8; Hervorhebung d. Verf.) zunächst das allgemeine optische Anforderungsprofil einer Sportart oder Bewegung bestimmt werden, indem folgende Gesichtspunkte beurteilt werden:

- statisch oder nicht-statisch (d. h. ist die optische Information in Bewegung?),
- dynamisch oder nicht-dynamisch (d. h. ist der Sportler nahezu unbewegt bei der Ausübung einer Sportart?),
- zeitliche Dauer der Bewegungsausführung (d. h. auch die Zeit für die optische Informationsverarbeitung),
- Objektgröße (d. h. wie nah kann ich mich dem Beobachtungsobjekt nähern, ohne die Bewegungsausführung zu beeinflussen?),
- Äußere Beobachtungseinflüsse (Kontrast, Hintergrundfarbe, Schatten usw.).

Für die Schlagwurfbewegung lässt sich ausführen, dass sich der Werfer vom Anlauf bis zur Wurfauslage dynamisch bewegt, in der Hauptphase des Abwurfs dann aber seine Anlaufgeschwindigkeit bis zum Abstoppen abbremst, so dass man durchaus von einer nicht-dynamischen Endposition sprechen kann. Dafür werden in der Wurfauslage aber die Teilbewegungen von Hüfte, Schulter und Wurfarm beim Abwurf explosiv ausgeführt, so dass man von einer durchgängig nicht-statischen Bewegung sprechen kann, bei der sich die Dynamik vom ganzen Körper auf einzelne Körperpartien überträgt. Um den zeitlichen Umfang der zu beurteilenden Bewegungsausführungen möglichst einheitlich zu gestalten, wird der Anlauf auf fünf Schritte reduziert. Somit dauert die Bewegung nur zwischen zwei und drei Sekunden, die Zeit für die Informationsverarbeitung mit anschließender qualitativer Beurteilung wird nicht begrenzt. Realistisch betrachtet verblasst aber der wahrgenommene Bewegungseindruck relativ schnell, so dass ein Urteil über die Bewegungsqualität nicht zu lange hinausgezögert werden sollte. Die

Größe der zu beurteilenden Objekte ist durch technische Entscheidungen festgelegt. Da die Bewegungsausführung ohne Schwenk gefilmt wird, bewegt sich der Werfer vom linken zum rechten Bildschirmrand. Da die Wiedergabe auf einem 86-Zoll großen Smart-TV erfolgt, ist dabei das Objekt groß genug zu sehen. Der Abstand von 3m zum Smart-TV wurde daher als optimal betrachtet, um den Abstand zum Werfer realitätsnah nachzubilden. Um die äußeren Beobachtungseinflüsse möglichst gering zu halten, ist es von Vorteil, die Wurfbewegungen der zu bewertenden Schülerinnen und Schüler in der Turnhalle vor einem stabilen Hintergrund (Hallenwand) und mit künstlicher Beleuchtung aufzunehmen. Bei Bewegungsausführungen im Freien sind die Aufnahmen entsprechend aufwändiger, um eine gute Beobachtbarkeit hinsichtlich Helligkeit, Kontrast und Farbigkeit der Aufnahmen zu gewährleisten und äußere Störungen (z. B. durch Bewegungen im Hintergrund) zu vermeiden. Aufgrund dieser intensiven Vorüberlegungen können jedoch für jede mögliche Aufnahmesituation die bestmöglichen Bedingungen realisiert werden.

2.2.2.2 Wahl des Aufmerksamkeitsfokus

In diesem Kapitel sollen Informationen zusammengefasst werden, die sich mit dem Aufmerksamkeitsfokus beschäftigen, d. h. es wird der Frage nachgegangen, wie eine intentionale Bewegungsbeobachtung grundsätzlich ablaufen kann und von welchen äußeren Entscheidungen hinsichtlich der Beobachtungsstrategie getroffen werden können, um die Komplexität der Bewegungsbeobachtung zu differenzieren. Ausgehen sollen die weiteren Erörterungen aber von den Erkenntnissen der Blickforschung.

2.2.2.2.1 Eye-Tracking als Indikator für die Beobachtungsqualität?

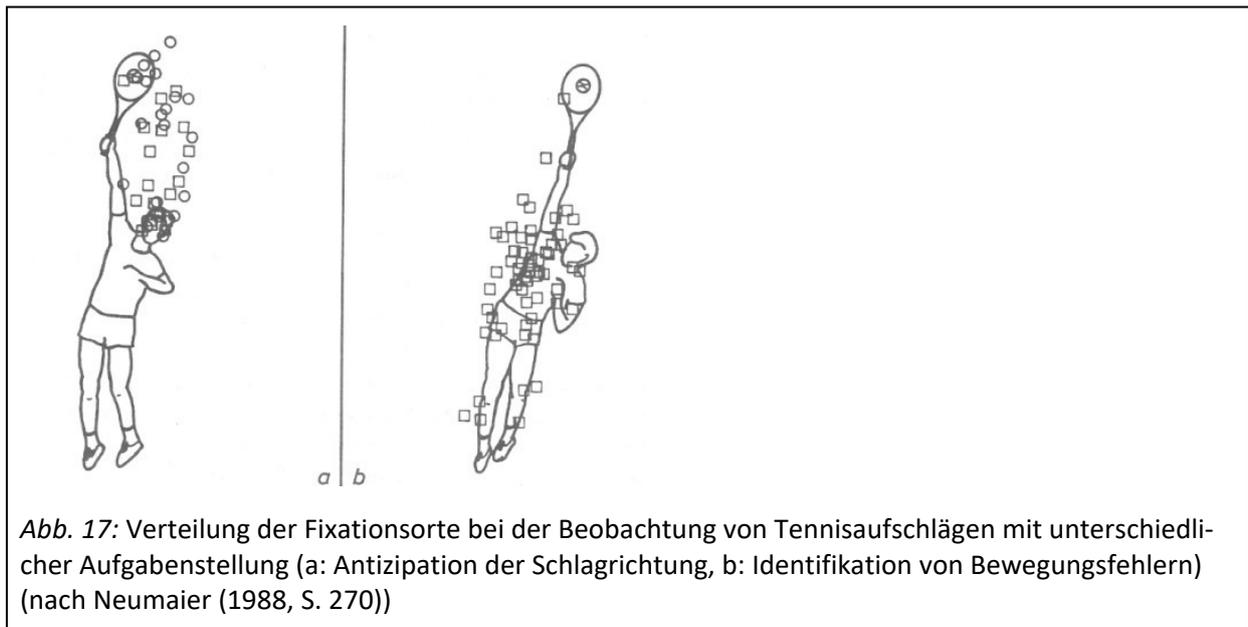
Geht man davon aus, dass der Fixationspunkt des Auges ein zuverlässiger Indikator dafür ist, auf welche Region oder welchen Körperbereich die visuelle Aufmerksamkeit gerichtet ist und welche visuellen Informationen folglich aufgenommen werden (Gabler, 2004, S. 53), so kann über das entsprechende Blickverhalten auch die Sehstrategie eines Beobachters ermittelt werden. Mittels Eye-Tracking-Systemen kann die Augenbewegung erfasst und die Fixationspräferenz aufgezeichnet werden, so dass unter anderem folgende, für die Bewegungsbeurteilung wichtige Merkmale des Blickverhaltens erfasst werden können:

- Fixationsorte
- Dauer von Einzelfixationen und durchschnittliche Fixationsdauer
- Zeitliche Anordnung und Reihenfolge der Fixationen
- Fixationshäufigkeiten bzw. Schwerpunkte in der Verteilung der Fixationen
- Anzahl der Sakkaden und Augenfolgebewegungen

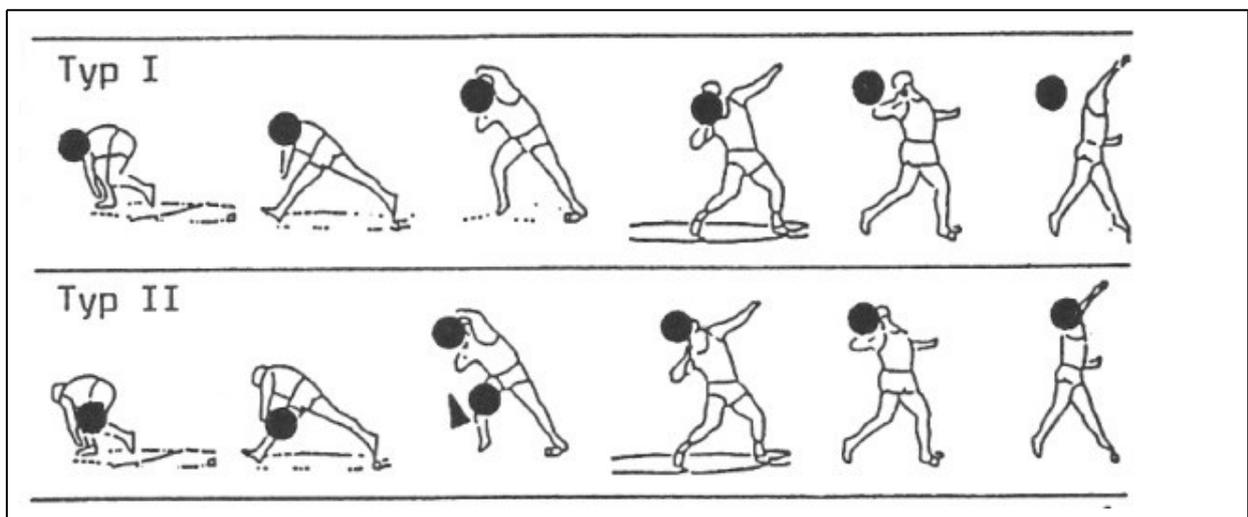
Einschränkend muss aber auch festgehalten werden, dass das Bestimmen des Fixationsortes noch keine Aussagen über das Ausmaß im peripheren Sehbereich zulässt. Oftmals wird der Blick gezielt an einer bestimmten, günstigen Stelle fixiert, von der aus der Gesamtverlauf der zu beobachtenden Bewegung über das periphere Sehen optimal erfasst bzw. überwacht werden kann (afferente Bewegungswahrnehmung). Ebenso kann es vorkommen, dass trotz Fixation keinerlei visuelle Information aufgenommen werden kann, weil es aufgrund zu großer Informationsdichte zu Unterbrechungsaktivitäten zugunsten der Verarbeitung bereits aufgenommener Informationen kommen kann. Somit ist das

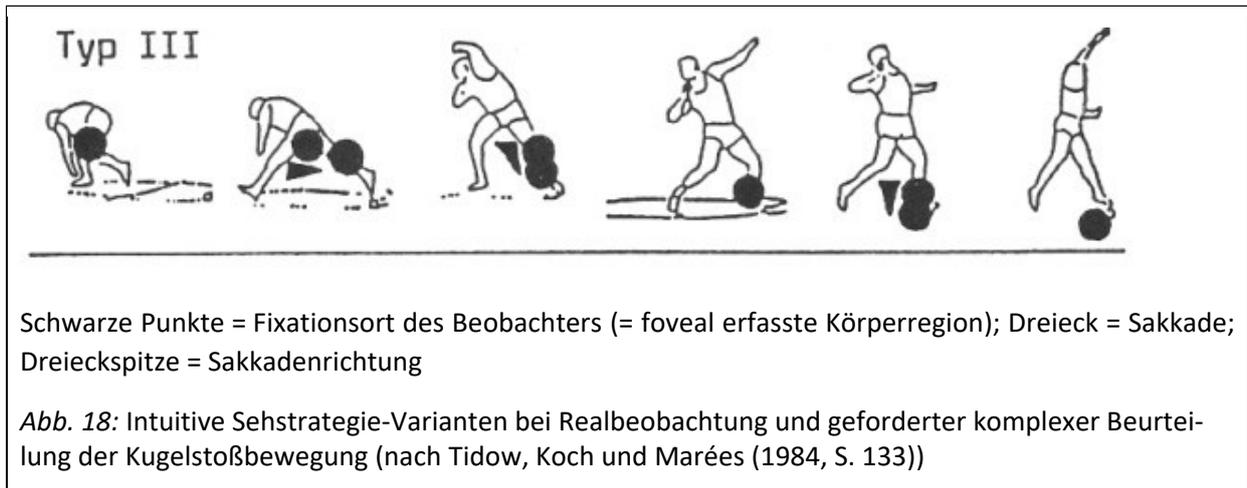
Feststellen einer Fixation nicht zwangsläufig mit dem Prozess der Informationsaufnahme und -verarbeitung gleichzusetzen. Ähnliche Einschränkungen gelten auch für die Interpretation der Fixationshäufigkeit und Fixationsdauer, von welchen nicht automatisch auf die visuelle Aufmerksamkeitslenkung geschlossen werden kann. Somit müssen mehrere Merkmale als integrative Bestandteile des komplexen Wahrnehmungsvorgangs verstanden werden, die nicht isoliert voneinander für Deutungen herangezogen werden sollen (Neumaier, 1988, S. 245–252). Diese Erkenntnisse stehen teilweise den Aussagen aus Kapitel 2.2.1 (dynamische Sehschärfe) inhaltlich entgegen, was aber nicht als Widerspruch zu bewerten ist, weil bei der Gesamtbeurteilung sportlicher Bewegungen auch das periphere Sehen eine Rolle spielt. Insofern darf durchaus infrage gestellt werden, ob mittels einer von Eyetracking-Systemen aufgezeichneten Aufmerksamkeitssteuerung Rückschlüsse auf die Qualität der Bewegungsbeurteilung gezogen werden können.

Studien über das Blickverhalten gab es vor allem in den 1980er Jahren. Sinnvoll sind solche technischen Aufzeichnungen auch deshalb, weil sich der Beobachter seiner eigenen Sehstrategie zumeist überhaupt nicht bewusst ist (Tidow, 1996b, S. 258–260). Es stellt sich also die Frage, welche Merkmale des Blickverhaltens unter welchen Bedingungen brauchbare Indikatoren für die visuelle Wahrnehmung im Sport darstellen. Das Blickverhalten ist dabei sehr stark abhängig von der jeweiligen spezifischen **Aufgabenstellung** sowie der Motivation des Betrachters, da der Blick stets nach den (subjektiv) relevanten Informationen für die Aufgabenlösung sucht. Das Fixationsverhalten ist demnach unterschiedlich, ob man eine Bewegung wie den Tennisaufschlag beobachtet, um die Schlagrichtung zu antizipieren oder um mögliche Bewegungsfehler zu erkennen (siehe Abb. 17). Die Aufgabenstellung kann sich auch dahingehend unterscheiden, dass entweder eine ganzheitlich-globale Bewertung der Bewegungsgüte im Sinne einer Gesamtbeurteilung abgegeben wird (in Kategorien wie gut/schlecht, gelungen/nicht-gelungen, fehlerfrei/fehlerhaft) oder eine detaillierte Aussage zu Einzelheiten der Bewegungsausführung gemacht werden soll (z. B. zur Körpervorlage beim Vollspannstoß oder zur Schlägerstellung bei Rückschlagspielen) (Neumaier, 1988, S. 329f.). Radford (1990, S. 3–7) beschreibt die Bewegungsbeobachtung analog als **bottom-up processing**, bei dem die Aufmerksamkeit passiv und unbewusst auf bestimmte Umweltreizen gerichtet und dann verengt wird, oder **top-down processing**, bei welchem der Beobachter gleich zu Beginn die zu beobachtenden Details bewusst auswählt oder den Fokus auf spezifische räumliche und zeitliche Bewegungselemente legt. Weil top-down processing willentlich vom Beobachter gesteuert wird, fließen in diesen Beobachtungsprozess wesentlich stärker die Vorerfahrungen des Beobachters ein (z. B. im Bezug auf das, was beobachtet wird). Dieses Priming erlaubt dem Beobachter zwar eine beschleunigte Informationsaufnahme und Merkmalsidentifizierung bei komplexen und schnellen Bewegungsabläufen, schränkt aber die Aufmerksamkeit derart auf die zu erwarteten Merkmalsausprägungen ein, dass auf unerwartete Bewegungsabläufe nicht angemessen reagiert werden kann.

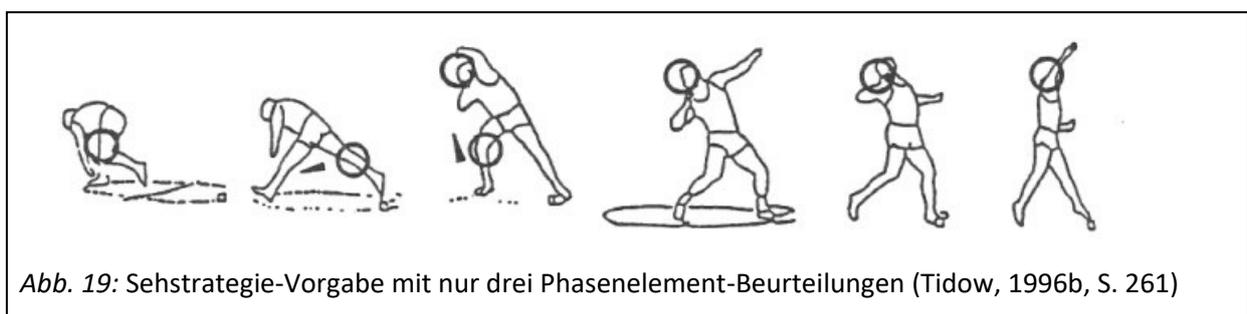


Um etwas konkretere Aussagen über das Blickverhalten bei der Bewegungsbeurteilung machen zu können, verglich Tidow (1996b, S. 258–261) die **Sehstrategien** von Leichtathletik-Trainern mit denen von Sportstudierenden und kam zu dem Ergebnis, dass drei verschiedene Vorgehensweisen bei der Beobachtung der O’Brien-Technik im Kugelstoßen zur Anwendung kamen. Entweder lag der Fokus schwerpunktmäßig auf der oberen Körperregion (Oberkörper und Wurfarm, Typ I in Abb. 18) oder der unteren (Beine und Hüfte; Typ III in Abb. 18). Bei der dritten Gruppe kam es zu einem Wechsel der Fixationsebene gemäß dem Bedeutungsgehalt der Phasen, indem zuerst die Beinebene, dann mittels einer Sakkade nach oben die Oberkörper-Stoßarm-Ebene fixiert wurde (Typ II in Abb. 18). Nach der Beobachtungsaufgabe mussten die Versuchspersonen die Bewegungsqualität anhand eines zwölf Elemente umfassenden Analysebogens beurteilen, wobei im Mittel ungefähr 20 Prozent der Bezugskriterien nicht ausgefüllt werden konnte. Überraschend ließ sich allerdings kein signifikanter Zusammenhang zwischen den Urteils-Aussagen und nicht-fixierten Körperregionen nachweisen, d. h. dass ein Kriterium nicht beurteilt werden kann, liegt nicht zwangsläufig an einer fehlenden Fixation der betreffenden Körperregion, was wiederum die Anfangs dieses Kapitels genannte These unterstützt, dass die Blickfixierung keinen geeigneten Operator für die Beurteilungsleistung darstellt.





In einer Untersuchung, ob der Analysebogen der Kugelstoßbewegung (siehe Abb. 29) in der ad-hoc-Beurteilung die Beobachtungskompetenz unterstützt, kam Tidow (1996b, S. 255–258) zu dem ernüchternden Ergebnis, dass nur 43 Prozent aller Bewertungen von Phaselementen voll mit einem nachträglich erstellten Filmurteil übereinstimmten, 7 Prozent keinerlei Übereinstimmung besaßen und 17 Prozent aller Phaselemente ausgelassen wurden. Die Darbietungscharakteristik und Komplexität dieses Bewegungsablaufs lässt also eine vollständige Detailbeurteilung nicht oder nur bedingt zu, was angesichts der 32 Einzelkriterien, welche der Analysebogen umfasst, zunächst nicht überrascht. Doch auch bei reduzierter Komplexität der Beobachtungsvorgaben (lediglich die drei Phaselemente Schwungbeinverhalten, Gleitbeinknie nach dem Aufsetzen und Ausstoßverhalten des Wurfarms sollten beobachtet werden, siehe Abb. 19) war ein Großteil der Beobachter nicht in der Lage, die Merkmale korrekt zu beurteilen und der Sehstrategievorgabe bis zum Schluss zu folgen. Die Blickbewegung vor allem beim Sakkadensprung vom Gleitbeinknie zum Stoßarm wurde von vielen zu spät gestartet oder verfehlte ihr Ziel mehr oder minder deutlich. Wenn die externe Sehstrategie jedoch auf Anhieb realisiert werden konnte, kam es auch zu sehr guten, korrekten Bewertungen des beobachteten Bewegungsablaufs. Tidow (1996b, S. 261) vermutet, dass Defizite im **antizipatorischen Timing** sowie ein partiell überfordertes visuell-dynamisches Leistungsvermögen dafür verantwortlich sein könnten. Umso wichtiger ist es, die zu beobachtende Bewegung genau zu kennen und sich eine Beobachtungsstrategie vorzubereiten, welcher man in der Beurteilungssituation folgen kann.



Petrakis (1993) verglich in der Sportart Tennis die Beobachtungsstrategien von Trainernovizen und erfahrenen Coaches anhand des Vorhandschlags und Aufschlags. Er fand dabei heraus, dass es in Bezug auf die Anzahl sowie Dauer der Augenfixierungen (number and duration of eye fixations) keinen signifikanten Unterschied zwischen beiden Gruppierungen gab, hingegen bezüglich des Ortes und der Bahn

der Fixierungen (scan patterns) schon Auffälligkeiten zu Tage traten. Die erfahrenen Trainer fokussierten kompakter auf die entscheidenden Körperregionen (Körperschwerpunkt beim Vorhandschlag und Oberkörper/Schläger beim Aufschlag), während die Novizen ihren Blick breiter auf mehrere Körperregionen schweifen ließen.

2.2.2.2 Beobachtungsstrategien

Anhand der eben beschriebenen Analyse der Blickaufzeichnungen wird schon ersichtlich, dass beim Beurteilen von Bewegungen intuitiv unterschiedliche Beobachtungsstrategien angewendet werden. Welche Körperregion wie lange fixiert wird, hängt maßgeblich vom Ziel der Bewegungsbeobachtung ab. Daher wird zunächst theoretisch geklärt, auf welche Arten eine Bewegungsbeobachtung ablaufen kann, um anschließend zu diskutieren, welche Beobachtungsstrategie für die Beurteilung der leichathletischen Schlagwurfbewegung sinnvoll erscheint und bei der Konstruktion des Beurteilungstests berücksichtigt werden soll.

Es gibt nach Knudson (2013, S. 102) vier mögliche Schwerpunkte, an denen sich eine scanning strategy orientieren kann: Die Beobachtung kann (1) entlang der einzelnen Bewegungsphasen erfolgen, (2) die wichtigsten Beobachtungsmerkmale umfassen, (3) sich vom allgemeinen Eindruck auf spezifischere Handlungen verengen oder (4) auf Gleichgewichtsvorgänge und -verläufe gerichtet sein.

Die häufigste Beobachtungsstrategie besteht darin, bestimmte critical features entsprechend ihrer Reihenfolge oder Phase im Bewegungsverlauf visuell zu erfassen (**observation by phases of movement**). Die Informationsdichte wird reduziert, indem der Fokus auf der Wahrnehmung der drei Phasen Vorbereitung (preparation), Ausführung (action) und Abschluss bzw. Übergang (follow-through) liegt.¹⁴ Dabei sollte auch geklärt werden, welche Phase zuerst beobachtet werden soll, weil dies trotz klarer zeitlicher Abfolge nicht immer chronologisch erfolgen muss. Beim Schlagwurf würde man sich zuerst sinnvollerweise auf die Hauptphase des Abwurfs (zeitlich gesehen von der Wurfauslage bis zum Verlassen des Wurfgeräts aus der Hand) konzentrieren, ehe man den Anlauf mit Dreischritt-Rhythmus (Vorbereitung) oder das Abfangen der Wurfbewegung (Endphase) beobachtet. Kenntnisse über Einzelbewegungen verschiedener Körperteile während der drei Bewegungsphasen sind die Voraussetzung dafür, dass die Teilbewegung dann auch während der Bewegungsausführung beobachtet werden können. Dabei kann die Bewegung des ausgewählten Körperteils (z. B. des Wurfarms beim Schlagwurf) kontinuierlich beobachtet werden, um Aussage über den Umfang der Ausholbewegung oder die Kopplung der Wurfarmbewegung treffen zu können. Beim nächsten Bewegungsvollzug des Athleten kann der Fokus dann auf eine andere Körperregion gelegt werden (z. B. auf die Beinaktionen beim Impuls- und Stemmschritt).

Logan und McKinney (1977, S. 213) empfehlen in einem *segmental approach* für eine möglichst umfassende Bewegungsbeobachtung beispielsweise acht Durchgänge, die nacheinander auf folgende Aspekte oder Körperteile gerichtet sein sollten: (1) Gesamtbewegung, (2) Becken und Brustkorb, (3) Standfläche oder Füße, (4) Kopf und Schultern, (5) Arm und Hand, (6) Knie und Hüfte, (7) Bewegungsabschluss und erneut die (8) Gesamtbewegung. Um sich bewusster auf die Beobachtung vorzubereiten, können je nach technischem Ausführungsniveau der Bewegung unterschiedliche Fragen an die acht Beobachtungsdurchgänge gestellt und beantwortet werden (siehe Tab. 3).

¹⁴ Logan und McKinney (1977, S. 224) ergänzen noch die Phasen *stance* (geht der Vorbereitungsphase voraus) und *recovery* (folgt der follow-through Phase).

Tab. 3: Fragen zur Bewegungsbeobachtung der acht Beobachtungsdurchgänge nach Logan und McKinney (1977, S. 213f.)

<p>I. Total Performance</p> <ul style="list-style-type: none"> A. What is the general timing on the part of the athlete for the total skill? B. What was the outcome of the performance?
<p>II. Pelvic Area and Rib Cage. The pelvic girdle is a relatively slow movement area and a good orientation point for analysis when compared to limb action</p> <ul style="list-style-type: none"> A. How high is the pelvic girdle from its base? B. What is the critical motion of the pelvic girdle? C. Does the critical path change during any phase of the movement? D. What is the extent and direction of the total pelvic action?
<p>III. Base of Support</p> <ul style="list-style-type: none"> A. How does the body weight shift at the start, during, and at the finish of action? B. What is the line of projection for efficient action? C. Does the line of projection during the performance change? D. In what direction(s) do the feet point prior to, during, and at the completion of the performance? E. Are the feet too far apart or too close together during the performance? F. What is the position of the feet immediately prior to the action?
<p>IV. Head and Shoulders</p> <ul style="list-style-type: none"> A. Is the head in the proper position at the start, during, and at the end of the action? B. Are both eyes focused on the objective? C. Which is the performer's preferred eye? D. Were the eyes open or closed during the crucial phase of the performance? E. Is there any abnormal movement at the start of the performance which would have an adverse effect on subsequent movements? F. In cases where an implement is used, are the eyes focused on the object to be hit at the moment of impact?
<p>V. Arm and Hand</p> <ul style="list-style-type: none"> A. In what direction do the arms move during the action? B. Are the arms too close or too far from the body for effective movement? C. Are the arms moving diagonally or in flexion-extension patterns? D. In cases where an implement is used, is the proper grip taken initially and held throughout the performance? E. In cases of impact, was the implement released at the point of impact? F. Was the total range of motion of both arms adequate? G. Are both arms rotated properly?
<p>VI. Knees and Hips</p> <ul style="list-style-type: none"> A. In what direction do the legs move during the action? B. Are the knee flexion angles consistent with the skill to be performed? C. Are there any rotational movements at the hips which might inhibit performance? D. Is there any indication of inadequate flexibility at the hip joints? E. If impact is an integral part of the skill, what are the knee and hip actions immediately prior to, during, and following impact? F. Are the legs moving diagonally or in the flexion-extension patterns?

VII. Follow-through and/or Recovery

- A. What is the extent, direction, and pattern of the followthrough?
- B. Is there any evidence that muscular tension impeded the follow-through? If so, in what body segments?
- C. If an implement was used, how did it react at the point of impact and during the follow-through?
- D. Was the follow-through continuous?
- E. Was the recovery related to subsequent movements?

VIII. Total Performance

- A. Was there an effective summation of internal forces?
- B. Was the objective of the skill met?

Unabhängig von der gewählten Beobachtungsstrategie empfehlen einige Autoren, den Beobachtungsplan schriftlich auszuarbeiten, wobei dies in unterschiedlicher Form z. B. als Bewertungsrubrik, Checkliste, Diagramm, Beurteilungsskala oder Beobachtungsbogen geschehen kann (Knudson, 2013, S. 106). Gerade die Wahl des Beobachtungsfokus soll nach Barrett (1979) vor der Beobachtung in einem Beobachtungsplan (scanning strategy) festgehalten werden, der beschreibt, welches Objekt zu welchem Zeitpunkt der Bewegungsausführung wie lang fokussiert werden soll. Dabei sind es nicht nur die critical features, die beobachtungsrelevant sind, sondern auch andere Variablen, die sich auf den Sportler oder die Beurteilungssituation beziehen (z. B. regelkonforme Bewegungsausführung beim Badmintonaufschlag, Umwelteinflüsse wie Windverhältnisse beim Beachvolleyball, unrythmische oder überflüssige Bewegungsausführungen und Ermüdungserscheinungen). Pinheiro und Cai (1999) bestätigen, dass Beobachter, welche einen Beobachtungsbogen verwendeten, sportliche Bewegungen besser diagnostizierten als solche ohne derartige Hilfsmittel. Daher soll auch im Rahmen dieser Studie ein schriftlich ausgearbeitetes Beobachtungsraster entwickelt werden, anhand dessen die Testpersonen ihre Beobachtung ausrichten und ihre Beurteilung vornehmen zu können (siehe Tab. 53 im Appendix).

Qualifizierende Beschreibungen (z. B. der Bewegungsrichtung oder einer Gelenkposition), welche für eine bestimmte Bewegung in Form einer Beobachtungsvorlage zusammengefasst werden, helfen dem Beobachter, sich auf die Beobachtungsaufgabe vorzubereiten. Besonders geeignet ist diese Sehstrategie für open motor skills, d. h. für Bewegungen, deren Verlauf durch externe Faktoren (z. B. Gegnerverhalten, Umwelteinflüsse) beeinflusst wird. Aber auch für die Beobachtung des Schlagwurfs kann ein solches observational model (siehe Tab. 4) sinnvoll sein, um die Beobachtungsmerkmale von Körperteilbewegungen in den einzelnen Bewegungsphasen zu vergegenwärtigen. Im Rahmen dieser Studie wird es für sinnvoll erachtet, ein solches Beobachtungsmodell visuell zu entwerfen, d. h. den Testpersonen wird anhand von Demonstrationsvideos erklärt, wie ein Bewegungsmerkmal bei einem idealen Schlagwurf ausgeführt werden soll. Ebenso soll darin thematisiert werden, mit welchen geringen oder starken Abweichungen zu rechnen ist, damit eine einheitliche Einschätzung der Bewegungsqualität von den verschiedenen Testpersonen vorgenommen werden kann (siehe Kapitel 4.5.2), die im Wesentlichen auf der gleichen Beobachtungsstrategie erfolgt.

Tab. 4: Beobachtungsvorlage für den Überkopfwurf mit Einteilung in drei Bewegungsphasen nach Gangstead und Beveridge (1984, S. 62)

Body components	TEMPORAL PHASING		
	Preparation	Action	Follow-through
Path of hub	Over base of support	Shift forward to target	Continue movement to target
Body weight	Over base of support	Shift forward to target	On front foot closest to target
Trunk action	Nonthrowing side to target	Rotate open to target	Follow arm to target
Head action	Face target	Eyes on target	Eyes on target
Leg action	Apart, weight on back leg	Step to target with closest leg	Bring back leg up to front leg
Arm action	Throwing arm extended back	Bring throwing arm forward	Throwing arm across body
Impact/release		Snap wrist	

Nimmt man den Standpunkt ein, dass eine spezielle Bewegungsfertigkeit im Ganzen mehr ist als die Summe der einzelnen Bewegungsteile, so muss der Fokus bei der Beobachtung zunächst auf der Gesamtbewegung liegen (Bottom-up attentional processing nach Radford (1990) oder Gestalt observational model nach Dunham (1994)). Wenn aufgrund dieses Gesamteindrucks beim Beobachter das Gefühl entsteht, dass im Bewegungsablauf etwas nicht stimmt, können etwaige Unstimmigkeiten durch einen genaueren Blick auf die einzelnen Bewegungsphasen oder die Bewegung einzelner Körperteile identifiziert werden. Wenn der Beobachter also aufgrund dieses ersten Gesamteindrucks eine Schwäche an einer bestimmten Stelle des Bewegungsablaufes (z. B. dem Bewegungsrhythmus oder dem Bewegungsfluss) erkennt, kann die Beobachtungsvorlage zur genaueren Identifikation eines möglichen Fehlerbildes zur Hand genommen werden. Zugleich können die in Tab. 4 verwendeten Kategorien aber auch für die Konstruktion eines offeneren Beobachtungsbogens dienen, in den man bestimmte Auffälligkeiten im Bewegungsablauf eintragen kann (siehe Abb. 113 im Appendix). Dieser Ansatz wird speziell bei der Beobachtungsaufgabe der Fehleranalyse (Teil D des Beurteilungstests, siehe Kapitel 4.5.3) verfolgt, bei der die Testperson nach der Betrachtung des gesamten Bewegungsablaufs diejenigen Bewegungsmerkmale identifizieren soll, welche stark vom Sollwert abweichen.

Eine Beobachtungsstrategie kann auch anhand der Wichtigkeit der identifizierten critical features entwickelt werden (Morrison (1994)). Ein Bewegungsmerkmal, das andere Aspekte des Bewegungsverhaltens beeinflusst (wie z. B. die Gleichgewichtsfähigkeit bei der Skatingtechnik im Langlauf), sollte daher bevorzugt beobachtet werden (**observation based on importance**). Erfahrene Trainer, die eine Bewegung genau kennen, haben meist eine sehr klare und praxiserprobte Meinung, welche critical features die meiste Aufmerksamkeit verdienen. Biomechanische Modelle können die Auswahl der wichtigen Bewegungsmerkmale zusätzlich objektiv absichern, indem die biomechanische Bedeutung der beobachteten Variable für das Bewegungsziel nachgewiesen wird (Knudson, 2013, S. 105). Diese Beobachtungsstrategie kommt im Beurteilungsteil E (synchrone Gesamtanalyse, siehe Kapitel 4.5.3) teilweise zur Anwendung, weil zwar alle Bewegungsmerkmale beobachtet werden sollen, jedoch die Testperson während der Bewegungsbeobachtung selbst entscheidet, welche Bewegungsmerkmale sie entsprechend ihrer Wichtigkeit bewusster wahrnimmt als andere.

Eine vierte Möglichkeit besteht vor allem bei Sportarten, in denen das Gleichgewicht ein leistungslimitierender Faktor ist, darin, der Gleichgewichtsfähigkeit und dem Bewegungsursprung eine besondere Aufmerksamkeit zu schenken. Die Unterstützungsfläche und die Bewegungsinitiierung aus den unteren Extremitäten (z. B. beim Skifahren) erscheinen beobachtungswürdig, weil sie die folgenden Bewegungsphasen entscheidend beeinflussen können. Sogar in Sportarten, in denen man es auf den ersten Blick nicht vermutet, kann eine Rückmeldung über die Gleichgewichtsfähigkeit für die Bewegungsausführung wichtiger sein als konkrete Anweisungen für Teilbewegungen von Körperteilen (z. B. beim Baseballwurf) (Knudson, 2013, S. 104f.). Auch beim Schlagwurf spielt das Gleichgewicht eine Rolle, denn die Stabilität in den unteren Extremitäten dient als Voraussetzung für eine gelungene Wurfausführung des Oberkörpers und Wurfarms. Ist der Stemmschritt zu kurz gesetzt oder fällt der gesamte Körper beim Werfen zur Seite, so verlässt der Körperschwerpunkt die Unterstützungsfläche der Beine und es kommt zu einem schwer zu kompensierenden Gleichgewichtsverlust. Dies kann schon während der Hauptphase der Wurfbewegung beobachtet werden, zeigt sich aber speziell in der Endphase der Bewegung, wenn es darum geht, den Körper nach der Wurfbewegung abzufangen und die Abwurflinie nicht zu übertreten. Insofern spielt diese auf das Gleichgewicht gerichtete Beobachtungsstrategie vor allem bei der Beurteilung des Merkmals Bewegungspräzision eine Rolle, das sich vor allem in der Phase des Abfangens der Wurfenergie (z. B. durch ein möglichst gestrecktes Stemmbein) zeigt (siehe Kapitel 4.5.1). Im Beurteilungsraster (siehe Appendix Tab. 53) drückt sich dieser Beobachtungsfokus beim Bewegungsmerkmal Präzision beispielsweise durch die Formulierung ‚Verlust des Gleichgewichts‘ als Kriterium für eine stark vom Sollwert abweichende Bewegungsausführung aus.

2.2.2.2.3 Variablen der Aufmerksamkeitsfokussierung

In diesem Kapitel wird überlegt, wie die Beurteilungskriterien beschaffen sein sollten, damit eine morphologische Beurteilung der Schlagwurfbewegung möglichst gut gelingen kann. Es stellt sich die Frage, in welcher Detailliertheit Bewegungsqualität überhaupt realistisch eingeschätzt werden kann, wenn man die Limitationen bedenkt, mit welchen die visuelle Informationsaufnahme erfolgt. Wie kleingliedrig soll die Beurteilungsskala angelegt sein, um das Urteilsvermögen der Testpersonen nicht zu überfordern? Die Erkenntnisse dieses Kapitels fließen in die Entscheidungen über die Konstruktion der Qualitätsstufen ein, mit deren Hilfe die Bewegungsqualität eingeschätzt und somit die Beurteilungsleistung operationalisiert werden soll (siehe Kapitel 4.5.2).

Grundsätzlich stellen ganzheitliche Bewegungsbeurteilungen, die sich auf allgemeine Bewertungskriterien wie Bewegungsharmonie, Eleganz oder Schönheit beziehen, geringere Anforderungen an den Beurteiler als die Wahrnehmung von Bewegungsdetails, welche eine differenzierte Bewegungsvorstellung mit der Bildung spezifischer Erwartungen im Wahrnehmungsvorgang voraussetzt. Ebenso nimmt die Komplexität der Beurteilungsaufgabe sehr stark zu, wenn über die allgemeine Einschätzung fehlerhaft/fehlerfrei hinaus auch noch die konkrete Abweichung vom korrekten Bewegungsablauf genannt werden sollte. Während die globalere Einschätzung auch von Beurteilern mit wenig Beurteilungserfahrung schnell gelernt werden kann, weil eine ganzheitliche Repräsentation der Bewegung mit wenigen kennzeichnenden Merkmalen als kognitives Vergleichsmaterial ausreicht (Baumann, 1986, S. 52f.), ist eine detaillierte Fehleranalyse nur für Beobachter möglich, die über einen differenzierten Sollwert verfügen.

Wenn der Beurteiler weiß, welche Fehlerbilder in welchen Phasen des Bewegungsablaufs auftauchen können, so nimmt die Wahrscheinlichkeit der Fehleridentifikation zu, wohingegen neue, unerwartete Sollwertabweichungen häufig übersehen und nicht richtig benannt werden können (Radford, 1990, S. 7). Ebenso spielt das Ausmaß der Abweichung vom Sollwert eine Rolle, denn je deutlicher der Fehler ausgeprägt ist, desto augenfälliger wird er auch für den Beobachter (Neumaier, 1988, S. 329–335). Die Vorinformationen und Vorerfahrungen, die ein Beobachter sich angeeignet hat, prägen also maßgeblich die **Erwartungshaltung** gegenüber der zu beobachtende Bewegung, welche ihrerseits die Sehstrategie in positiver wie negativer Weise beeinflussen kann. Daher gehört zu einer umfassenden Schulung der Beurteilungsfähigkeit auch die Vermittlung von besonders häufig auftretenden Fehlerbildern, um das Wahrnehmungssystem zu entlasten. Hofmann und Armstrong (1975) sowie Gangstead (1984) fanden diesbezüglich heraus, dass Beobachter, deren Bewegungsvorstellung mittels korrekter und fehlerhaft ausgeführten Bewegungsmustern geschult wurde, bei der Fehleridentifikation bessere Leistungen zeigten als solche, die nur mit korrekten Bewegungsausführungen konfrontiert wurden.¹⁵

Bei der Detailanalyse spielt es noch eine Rolle, ob das Objekt zeitraumbezogen in einer ausgeführten **Aktion** (z. B. während des Schmetterschlags) oder zeitpunktbezogen bezüglich flüchtig oder statisch eingenommener **Positionen** (z. B. Gelenkwinkel, Positionen mehrerer Körperteile zueinander) beurteilt werden soll. Obwohl für die Wahrnehmung von Aktionen in der Regel mehr Zeit zur Verfügung steht als für das Erkennen von Positionen, kann man daraus nicht generell schließen, dass erstere besser wahrnehmbar und beurteilbar sind. Tidow (1993, S. 315) weist bei der Beobachtung positionsbezogener Merkmale während der Kugelstoßbewegung Wahrnehmungsschwächen nach, ebenso wie Ballreich (1981, S. 517) berichtet, dass beim Hochsprung ausgeführte Aktionen bezüglich der Zeit- und Geschwindigkeitsverläufe von Körperteilen oder des Absprungkrafteinsatzes zuverlässiger beobachtet werden können als Körperpositionen (z. B. die maximale Kniebeugung beim Absprung). Neumaier (1988, S. 336–338) hingegen verweist auf Studien, bei denen Positionen genauer eingeschätzt werden als Aktionsmerkmale (z. B. beim Weitsprung in der Leichtathletik oder beim Flickflack im Gerätturnen). Die visuelle Wahrnehmbarkeit einer Aktion bzw. Position hänge aber nicht nur von der verfügbaren Wahrnehmungszeit ab, sondern auch von der Anzahl der zu berücksichtigenden Beobachtungseinheiten, der zeitlichen Dichte, in der die einzelnen Beobachtungseinheiten aufeinander folgen, sowie der Reihenfolge, d. h. Stellung der einzelnen Beobachtungseinheiten zueinander. Generalisierende Aussagen sind also kaum möglich, sondern vielmehr nur für die jeweilige konkrete Beurteilungsaufgabe und die Beobachtungsbedingungen gültig.

Eine Rolle für die Sehstrategie bei der Bewegungsbeurteilung spielt auch die Tatsache, ob **qualitative oder quantitative Bewegungsmerkmale** eingeschätzt werden sollen. Grundsätzlich sind qualitative Merkmale (wie Bewegungstempo, Bewegungsfluss) leichter zu beobachten, weil ihre Wahrnehmung keine besonders spezifischen und präzisen Vorstellungen vom Bewegungsablauf voraussetzt. Diese unterliegen aber stärker den subjektiven Vorstellungen des Beurteilers, wodurch eine geringere Beurteilungsobjektivität zustande kommt. Bei quantitativer Bewegungsbeurteilung besteht die Schwierigkeit darin, solche mit Messinstrumenten sehr genau bestimmbare dynamische und kinematische Parameter (wie z. B. Gelenkwinkel, Anlaufgeschwindigkeit, Absprungkrafteinsatz) mit dem Auge möglichst exakt zu erfassen. In oben bereits genannter Untersuchungen fand Ballreich (1981, S. 517) her-

¹⁵ Die Forschungslage hierzu ist insgesamt indifferent. Morrison (1994, p.169) fand in seiner Untersuchung heraus, dass es keine Auswirkung auf die Bewegungsbeobachtungsfähigkeit einer Lehrperson hat, ob mit guten und schlechten Bewegungsausführungen konfrontiert wurde oder ausschließlich mit guten (ebenso auch Morrison & Reeve (1989)).

aus, dass Zeit- und Geschwindigkeitsschätzungen in hohem Maße und Schätzungen des Absprungkraftereinsatzes relativ gut mit den biomechanischen Messwerten übereinstimmten, wohingegen Körpergelenkwinkel nur schlecht visuell erfassbar waren. Insgesamt ist die Fähigkeit zur quantitativen Bewegungsbeurteilung eher als mittelmäßig bis schwach einzustufen. Ob die Beurteilung von räumlichen gegenüber zeitlichen oder von kinematischen gegenüber dynamischen Merkmalen generell höhere Beobachtungsanforderungen stellt, lässt sich nicht abschließend beantworten (Neumaier, 1988, S. 339–348).

Grundsätzlich beeinflussen die spezifische visuelle Szene (z. B. Bewegungen am Gesichtsfeldrand, Helligkeits- und Kontrastverhältnisse) und die Beschaffenheit des **Beobachtungsgegenstands** (z. B. Objektgröße, auffällige Trikotfarbe) das Blickverhalten unbewusst (Neumaier, 1988, S. 268–275). Da die Bewegungsbeispiele als Filmmaterial präsentiert werden, müssen diese Rahmenbedingungen berücksichtigt werden. So kann beispielsweise auf einen stabilen Hintergrund ohne Störbewegungen geachtet werden (z. B. eine einfarbige Hallenwand), bei dem auch der Kontrast zu den Werferinnen und Werfern entsprechend hoch ist.

Zyklische Bewegungen (z. B. Lauf- oder Schwimmbewegungen) ermöglichen es dem Beobachter, bestimmte Details in zeitlich kurzem Abstand mehrmals hintereinander zu betrachten, wohingegen bei **azyklischen** Bewegungen (z. B. Wurfbewegungen) eine hohe Informationsdichte bei sehr kurzer Wahrnehmungszeit verarbeitet werden muss, was vor allem die Fehleridentifikation bei komplexen Bewegungsabläufen erschwert. Jedoch ist der zeitliche Phasenverlauf eines Bewegungsmerkmals oder der Gesamtbewegung bei zyklischen Bewegungen schwieriger zu beobachten, weil die Endphase eines Bewegungszyklus kontinuierlich in die Vorbereitungsphase des nächsten Zyklus übergeht (Gangstead & Beveridge, 1984, S. 62f.). Die Komplexität einer Beurteilungsaufgabe ist auch davon abhängig, ob verschiedene Beobachtungseinheiten zeitlich nacheinander (**sukzessiv**) oder **simultan** ablaufen. Zweites setzt derart hohe Erkennensleistungen in der entsprechenden Bewegungsphase voraus, dass in der vorhergehenden und folgenden Phase deutlich weniger Fehler identifiziert werden (Positionseffekte). Gleiches trifft auf Bewegungsphasen mit einer hohen Fehlerdichte, kurzer Darbietungszeit oder hoher Winkelgeschwindigkeit zu. Außerdem verschlechtert die Aufgabe, Simultanbeurteilungen vornehmen zu müssen, die Beurteilbarkeit der einzelnen Beobachtungseinheiten. Die Aufmerksamkeit wird dabei entweder auf einen bestimmten Fixationsort konzentriert (die Simultaneinheiten werden dann mit großer Wahrscheinlichkeit übersehen) oder sie wird auf mehrere Beobachtungsmerkmale aufgeteilt (die Beurteilungsgüte der einzelnen Beobachtungseinheiten sinkt). Tidow (1983, S. 255) spricht in dieser Beurteilungssituation von einem Selektionskonflikt, der nur unbefriedigend gelöst werden kann. Ein begrenzender Faktor ist dabei die Wahrnehmungszeit, die laut Tidow (1983, S. 315f.) mindestens 200 ms für jedes zu beurteilende Detail beträgt. Bedenkt man die Tatsache, dass für das Identifizieren eines Bewegungsfehlers das aufgenommene visuelle Abbild auch noch mit dem gespeicherten Sollwert verglichen werden muss, so erscheinen selbst die von Schubert und Zehl (1984, S. 246) angenommenen 250 ms nicht ausreichend. Bei Einsatz sakkadischer Augenbewegungen wird der zeitliche Abstand zwischen sukzessiv zu berücksichtigenden Beobachtungseinheiten aufgrund von Wahrnehmungssuppression sogar noch größer. Da viele azyklische Bewegungsabläufe meist nicht länger als ca. 1 s dauern, erscheinen demnach höchstens 2–3 detaillierte Beobachtungseinheiten sinnvoll (Neumaier, 1988, S. 353–367).

Diese Feststellungen müssen auch bei der Beurteilung der Schlagwurfbewegung berücksichtigt werden, die mit fünf Anlaufschritten ausgeführt nach ungefähr zwei Sekunden beendet ist und in der eigentlichen Hauptphase der Wurfbewegung nur wenige Zehntel Sekunden dauert. Demnach muss bei

der Konstruktion der Beurteilungsaufgaben darauf geachtet werden, dass durch die Anzahl der zu beobachtenden Bewegungsmerkmale keine Überforderung entsteht. Es ist ausreichend, zunächst nur ein Bewegungsmerkmal zu beurteilen, ehe die Komplexität über die zunehmende Anzahl erhöht wird. Außerdem ist zu berücksichtigen, dass in der Hauptphase der Wurfbewegung viele Teilbewegungen simultan ablaufen und damit sehr schwer gleichzeitig zu beurteilen sind. Es ist damit zu rechnen, dass beispielsweise die beiden Merkmale Bewegungsrhythmus und Bewegungskopplung aufgrund ihrer Sukzessivität leichter zu beurteilen sind als die direkt ineinander übergehenden Merkmale Bewegungsumfang und Bewegungskopplung. Gerade bei komplexen Beurteilungsaufgaben wie der Fehleranalyse erhöht sich die Anforderung der Bewegungsbeobachtung nochmal deutlich, weil eine simultane Wahrnehmung aller Bewegungsmerkmale erforderlich ist, um die stark abweichenden Bewegungsteile zu identifizieren.

Da die Zielgruppe dieser Untersuchung überwiegend unerfahrene Beobachter leichtathletischer Bewegungen umfassen wird, soll der ganzheitlichen qualitativen Bewegungsbeurteilung (mittels Beurteilungskriterien wie Bewegungsharmonie, Bewegungsumfang oder Bewegungsrhythmus) der Vorrang eingeräumt werden gegenüber der Bewertung allzu detaillierter (teilweise auch quantitativer) Technikmerkmale. Um diese globalen Bewegungsmerkmale beurteilen zu können, müssen zwar auch konkrete Positionen und Aktionen von Körperteilen wahrgenommen werden (z. B. spielt für die Einschätzung des Bewegungsumfangs bei der Ausholbewegung der Gelenkwinkel des Wurfarmellbogens eine Rolle oder für die Einschätzung des Bewegungsrhythmus die Beinaktionen beim Impuls- und Stemschritt), jedoch muss darüber keine detaillierte oder quantitative Rückmeldung erfolgen. Auf diese Weise ist es auch möglich, eine passende Einschätzung des Bewegungsmerkmals zu treffen, ohne genau begründen zu müssen, nach welchen Technikriterien die Einschätzung erfolgt ist.

Auch die Beurteilungsskala soll dabei so einfach wie möglich gestaltet werden, wenngleich eine dichotome Einschätzung in fehlerfrei oder fehlerhaft zu undifferenziert erscheint, weil das Ausmaß der Abweichung von der idealen Bewegungsausführung hier unberücksichtigt bleibt. Ein guter Beurteiler soll in der Lage sein, zwischen einer geringen und einer starken Abweichung vom Sollwert zu unterscheiden. Dies sollte bei der späteren Skalenbildung berücksichtigt werden. Da außerdem beim Beurteilen starke Abweichungen leichter einzuschätzen sind, sollen vor allem bei schwierigeren Beurteilungsaufgaben (wie z. B. der Fehleranalyse) die zu beurteilenden Bewegungsbeispiele idealerweise klar erkennbare Fehler zeigen.

2.2.2.2.4 Zusammenfassung

Es lässt sich zusammenfassend feststellen, dass die Sehstrategie sehr große Auswirkungen auf die Beurteilungsleistung hat. Wenn es gelingt, den Bereich des scharfen Sehens zum richtigen Zeitpunkt auf beobachtungsrelevante Körperregionen zu richten, so ist die Wahrscheinlichkeit groß, dass die Qualität eines Bewegungsmerkmals oder Bewegungsdetails auch richtig eingeschätzt werden kann. Wann auf welche Punkte der Bewegung geachtet werden muss, hängt entscheidend vom Wissen des Beurteilenden über den räumlich-zeitlichen Bewegungsverlauf sowie von seiner Erwartungshaltung bezüglich des Auftretens von Bewegungsabweichungen ab. Jedoch kann eine zu starke Fixierung auf das Beobachten bestimmter Bewegungsmerkmale oder Abweichungen dazu führen, dass Auffälligkeiten zu anderen Zeitpunkten der Bewegung oder in anderen Körperregionen nicht bemerkt werden. Wenn beispielsweise die Erwartung von Abweichungen beim Beobachtenden sehr stark in der Hauptphase der

Bewegung liegt, so kann es sein, dass deutlich auffällige Fehlerbilder in der Vorbereitungs- oder Endphase nicht bemerkt werden. Die Sehstrategie kann aufgrund der aufwändigen Messung in dieser Studie nicht als Variable der Beurteilungsleistung erfasst werden, wobei ohnehin festgehalten werden muss, dass allein die Aufzeichnung der Augenfolgebewegung und der Fixationsorte noch keine Rückschlüsse auf die Informationsaufnahme zulassen und somit nicht unbedingt als Indikator für die Beobachtungsgüte dienen können. Trotzdem konnten in diesem Kapitel wichtige Erkenntnisse über den Beobachtungs- und Beurteilungsprozess von Bewegungen gewonnen werden, die für die Testkonstruktion in dieser Studie von Belang sind. So kann beispielsweise aufgrund dieses Wissens über die Sehstrategien bei der Konstruktion des Beurteilungstests (siehe Kapitel 4.5) darauf geachtet werden, die Komplexität der Beurteilungsaufgaben schrittweise zu steigern. Ebenso können obige Ausführungen als Rechtfertigung dienen, die Bewegungsbeurteilung eher ganzheitlich an den Meinel'schen Bewegungsmerkmalen auszurichten als eine Engführung auf gewisse Beobachtungsdetails vorzunehmen, denn dies wird unerfahrenen Beobachtern, wie sie in dieser Studie vorliegen, besser gerecht und entspricht auch eher ihrer späteren Beurteilungstätigkeit im Schulsport.

2.2.2.3 Gestaltung der Beobachtungssituation und Beobachterposition

Da die zu beurteilenden Schlagwurfbewegungen nicht in einer echten, sondern in einer gefilmten Beobachtungssituation präsentiert werden, ist es wichtig, sich Gedanken über die Gestaltung dieser Beobachtungssituation zu machen. Die Schlagwurfbewegung ist zu den closed motor skills zu zählen, welche unter relativ stabilen Umweltbedingungen ausgeführt werden, da der Sportler selbst bestimmt, wann er die Bewegung beginnt, und er nur sehr selten auf äußere Störeinflüsse reagieren muss. Dies erleichtert das Arrangieren von Beobachtungssituationen, die für spätere Beurteilungsaufgaben gefilmt werden und grundsätzlich so gut kontrolliert werden sollten, dass die Bewegungsausführung des Sportlers so realistisch wie möglich ablaufen kann, ohne die Effektivität der Beobachtung allzu stark zu vermindern oder relevante Bewegungstechnikprobleme zu minimieren oder sogar zu verstärken. Allein schon die Tatsache, dass ein Beobachter zuschaut, hat je nach Sportler mehr oder weniger starke Auswirkungen auf seine Bewegungsausführung. Ebenso sollte die Beobachtungssituation nicht zu künstlich arrangiert werden. Relativ geschlossene Beobachtungsbedingungen sind vor allem auch gefordert bei systematischer Beobachtung oder beim Üben noch sehr instabiler Bewegungsmuster. Auch der Hintergrund, vor dem ein Sportler eine Bewegung ausführt, sollte idealerweise stabil sein (z. B. vor einer Wand), damit der Beobachter nicht durch Hintergrundaktionen abgelenkt wird. Horizontale oder vertikale Referenzlinien können helfen, Bewegungsrichtungen und Winkel besser einzuschätzen (Knudson, 2013, S. 106f.).

Außerdem kann der Blickwinkel des Beobachters (vantage points for observation) je nach Bewegung schon vorentscheidend sein über die Qualität der Beobachtung. Meist wird der Blick rechtwinklig auf die Bewegung (Seitperspektive) bevorzugt, weil man so die meisten Bewegungsmerkmale gezielt beobachten kann. Doch insgesamt empfiehlt sich aufgrund der dreidimensionalen Ausrichtung der meisten menschlichen Bewegungen, auch andere Beobachtungspositionen (z. B. die Frontalperspektive) einzunehmen, um je nach Bewegungsrichtung einen unverzerrten Blick auf die jeweilige Handlung zu haben. Die Entfernung des Beobachters vom Sportler ist ebenfalls ein entscheidendes Kriterium, weil diese auch die Augenbewegung des Beobachters beeinflusst, wenn die Bewegung mitverfolgt wird. Die Entfernung sollte daher so groß wie möglich sein, weil dadurch die Augenbewegung reduziert wird, dabei aber gleichzeitig noch die Beobachtung von Bewegungsdetails zulassen. Hay und Reid (1982)

empfehlen 10 bis 15 Meter Abstand für Bewegungen, die eine geringe Grundgeschwindigkeit haben oder nur geringe Bewegungsumfänge haben, während schnelle und ausladende (zumeist auch zyklische) Bewegungen von 20 bis 40 Metern aus beobachtet werden sollen. Beobachterdistanzen über 10 Meter sind aber manchmal schon aus räumlichen Gründen (z. B. in kleinen Hallen) nicht möglich und auch nicht immer ideal, wenn noch bewegungsbegleitende auditive, kinästhetische oder taktile Informationen (z. B. durch Hilfestellung für den Sportler) wahrgenommen oder dem Athleten direkt Rückmeldung gegeben werden soll. Daher empfiehlt Knudson (2013, S. 108) eine Minimaldistanz von 5 bis 10 Meter für die meisten menschlichen Bewegungen. Auch die Möglichkeit der Informationsaufnahme über peripheres Sehen sollte berücksichtigt werden bei der Wahl der Beobachterposition, denn das Nachvollziehen einer beobachteten Bewegung kann erschwert werden, wenn das periphere Blickfeld durch Kleidungsstücke (z. B. durch Kopfbedeckungen), Objekte oder andere Personen eingeschränkt wird. Besonders die Frage nach dem ‚Wohin‘ (d. h. der Richtung) einer Bewegung kann über peripheres Sehen klarer eingeschätzt werden und liefert ein besseres Verstehen der Bewegungsbeziehung verschiedener Körperteile zueinander.

Diese Bedingungen müssen beim Filmen von Schlagwurfbewegungen berücksichtigt werden, die für den späteren Beurteilungstest verwendet werden sollen. Da die Schülerinnen und Schüler in der Jahrgangsstufe 7 noch relativ instabile Bewegungsausführungen zeigen, wird versucht, die Beobachtungssituation möglichst einfach zu halten. Es wird ein Schlagwurf aus fünf Anlaufschritten verlangt, wobei die Startmarkierung und die Abwurflinie, die nicht übertreten werden sollte, vorgegeben werden. Die Schülerinnen und Schüler können selbst wählen, ob sie den Wurfarm schon rückgeführt haben oder die Rückführung während des Anlaufs ausführen. Es wurden keine weiteren technischen Vorgaben gemacht und alle Schülerinnen und Schüler führten zwei Wurfbewegungen nacheinander aus. Die Kameras werden so positioniert, dass sich die Übenden nicht gestört fühlen, die Gesamtbewegung aber ohne Kameraschwenk aufgenommen werden kann. Dies entspricht der Forderung aus der Beobachtungsstrategie, einen so großen Abstand zum Sportler zu wählen, dass die gesamte Bewegung nur mit Blickverfolgung und ohne Kopfdrehung beobachtet werden kann, um einen potentiellen Informationsverlust durch unkoordinierte Blicknachfolgebewegungen zu verhindern. Es wird dabei ausschließlich die Seitperspektive bei der Filmaufnahme gewählt, weil aus dieser die meisten Bewegungsmerkmale beobachtet werden können und durch die frontale Beobachterposition ein ungestörtes Ausüben der Wurfbewegung nicht mehr möglich wäre. Über diese theoretische Beschäftigung mit der Beobachtungssituation soll sichergestellt werden, dass die gefilmten Bewegungsausführungen ein möglichst realistisches Abbild der im Schulsport vorzufindenden Unterrichtsrealität darstellen und trotz des Perspektivverlusts gefilmter Aufnahmen die festgelegten Bewegungsmerkmale möglichst gut beobachtet werden können.

2.2.2.4 Beobachtungswiederholung

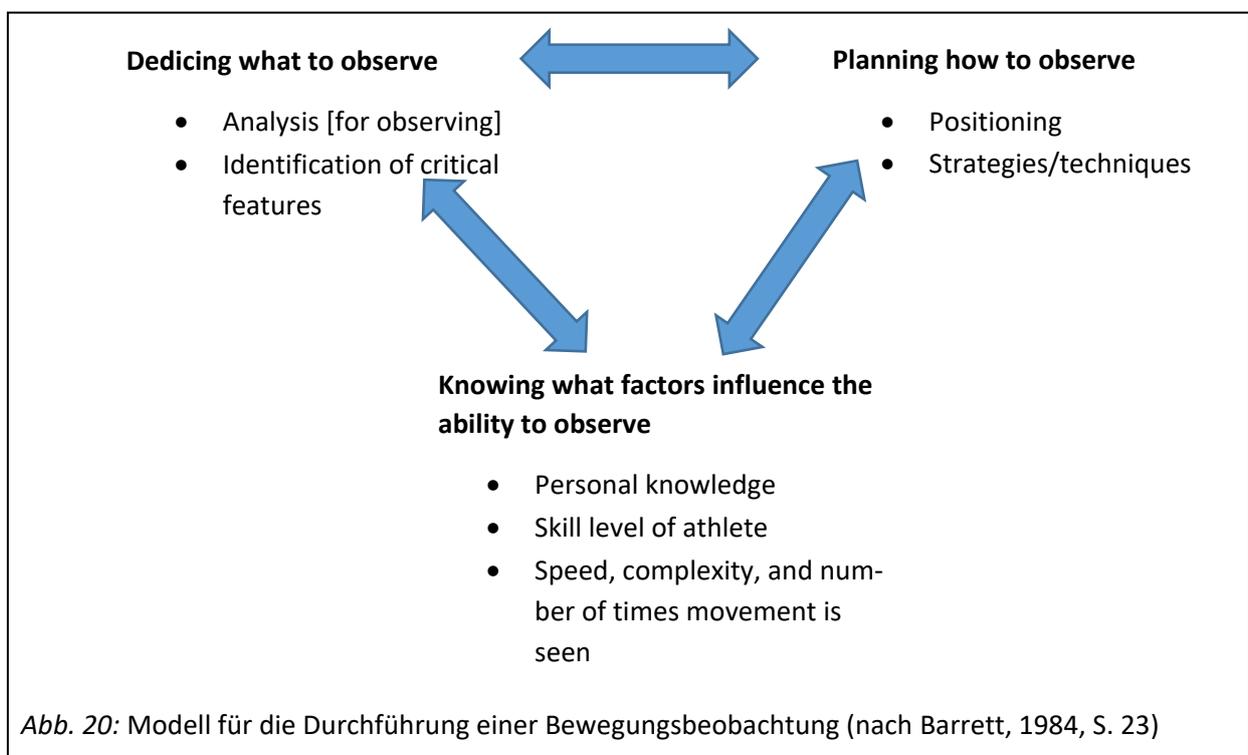
Es muss bei der Messung der Beurteilungsleistung vorher festgelegt werden, wie oft die zu beurteilende Bewegung beobachtet werden darf. Ein geschulter Beobachter weiß aus Erfahrung, wie viele Beobachtungsdurchgänge er benötigt, um trotz begrenzter optischer Wahrnehmungsfähigkeit (siehe Kapitel 2.2.1) genügend Informationen über die Bewegungsausführung eines Sportlers zu sammeln und entsprechende Techniknotenpunkte richtig zu beurteilen. Auch hier kommt man in manchen Beurteilungssituationen jedoch an Grenzen, weil beispielsweise im Schulsport nicht die Zeit bleibt, jeden Schüler entsprechend oft systematisch zu beobachten, bis alle Bewegungsphasen visuell erschlossen

sind. Beim Benoten einer Technikdemonstration muss die Lehrkraft zudem in der Lage sein, einen Bewegungsablauf bei nur einem Beobachtungsdurchgang differenziert zu bewerten. Speziell bei der Technikrückmeldung muss ein Kompromiss gefunden werden zwischen der optimalen Anzahl der Beobachtungsdurchgänge und der tatsächlich zur Verfügung stehenden Beobachtungszeit. Wie viele Versuche benötigt werden, hängt zum einen von der individuellen Beobachtungsfähigkeit, dem Ziel der Beobachtung oder der Komplexität der Bewegung ab. So reichen beispielsweise nach Clark, Stamm und Urquia (1979, S. 256) für Gleichgewichtsaufgaben von Kindern sechs Durchgänge zur Leistungsbeurteilung, während Logan und McKinney (1977, S. 213) mindestens acht empfehlen, bevor Rückmeldung gegeben werden sollte, und Hay und Reid (1982) sogar 15 Beobachtungsdurchgänge vorschlagen. Morrison und Reeve (1992) hingegen betonen, dass eine konsistente Technikrückmeldung der meisten sportlichen Bewegungen mit fünf Wiederholungen gut möglich ist. Obwohl einfache und langsam ablaufende Bewegungen auch nach einem Beobachtungsdurchgang beurteilt werden können, sollen trotzdem mehrere Versuche in die Bewertung und Intervention miteinfließen, um auch Aussagen über die Zuverlässigkeit der Bewegungsausführung treffen zu können. Daher reichen bei geübten Sportlern weniger Beobachtungsversuche aus als bei ungeübten, deren Bewegungskonstanz noch sehr gering ausgebildet ist (Knudson, 2013, S. 109f.).

Im Extremfall muss eine Lehrkraft aber beispielsweise bei der schulischen Leistungsbeurteilung in der Lage sein, aufgrund einer einzigen Bewegungsdemonstration eine entsprechend gut begründete Note zu bilden. Daher ist für die in dieser Studie angewandte Leistungsmessung nur ein Beobachtungsdurchgang erlaubt, auch wenn dies in den seltensten Fällen ausreicht, um einem Sportler oder Schüler eine ausreichende Bewegungsrückmeldung zu geben. Es kann angenommen werden, dass über diese hohe Anforderung die Messung der Beurteilungsleistung zwischen den Testpersonen differenzierter möglich ist, weil bei mehreren Beobachtungsdurchgängen nicht mehr so gut zwischen ‚guten‘ und ‚weniger guten‘ Beurteilern unterschieden werden kann.

2.3 Einflussfaktoren auf die Bewegungsbeurteilung

Nachdem gerade geklärt wurde, welche Einflussfaktoren bei der Bewegungsbeobachtung eine Rolle spielen, soll nun die Bewegungsbeurteilung näher in den Blick genommen werden. In einem hypothetischen Modell (siehe Abb. 20) weist Barrett (1984, S. 23) darauf hin, dass zwei Prozesse von entscheidender Bedeutung sind: (a) entscheiden, was beobachtet werden soll (siehe Kapitel 2.3.2) und (b) planen, wie beobachtet werden kann (siehe Kapitel 2.2.2). Das Wissen um Faktoren, welche die Beobachtungsfähigkeit (ability to observe) beeinflussen, steht nach diesem Modell ebenfalls in Wechselwirkung mit den beiden genannten Prozessen. So wird beispielsweise das eigene Bewegungswissen als möglicher Einflussfaktor auf die Beurteilungsleistung an späterer Stelle untersucht (siehe Kapitel 4.2), das Könnensniveau der in dieser Studie zu beurteilenden Athleten bedacht (siehe Kapitel 2.3.2.4) oder die Komplexität der Zielbewegung des leichtathletischen Schlagwurfs anhand verschiedener Bewegungsmerkmale behandelt (siehe Kapitel 4.5.1).



Die neue Herausforderung beim Beurteilungsprozess liegt also darin, einen beobachteten Bewegungsablauf (Istwert) mit der idealen Bewegungsausführung (Sollwert) zu vergleichen und anhand vorgegebener Kriterien zu bewerten. Wie schon zu Beginn von Kapitel 2 erwähnt, ist diese Beurteilungskompetenz nach Tidow (1993) abhängig von drei Faktoren, die in den folgenden Unterkapiteln näher erläutert werden:

1. Interne Sollwertpräzision und Fähigkeit zum Analogieschluss (siehe Kapitel 2.3.2)
2. Eigene Bewegungserfahrungen des Beobachters (siehe Kapitel 2.3.3)

Die theoretische Beschäftigung mit der internen Sollwertpräzision und der Fähigkeit zum Analogieschluss dient dazu, auf der Grundlage wissenschaftlicher Erkenntnisse eine möglichst präzise Bewegungsbeschreibung der leichtathletischen Schlagwurfbewegung anzufertigen, die nicht nur die räumlich-zeitliche Abfolge von Teilbewegungen umfasst, sondern auch biomechanische Aspekte, Fehlerwis-

sen und visuelle Schlüsselpositionen in sich vereint (siehe Kapitel 4.2). Bevor dieser forschungspraktische Schritt der Beschäftigung mit der Sollwertpräzision erfolgt (siehe Kapitel 2.3.2), soll aber eine theoretische Auseinandersetzung mit dem Begriff der Bewegungsvorstellung erfolgen, welche von Hartmann (1999) als ein die Bewegungsbeurteilung zentral beeinflussender Aspekt explizit genannt wird (siehe Abb. 10).

2.3.1 Bewegungsvorstellung

Der Bewegungsablauf stellt den gemeinsamen Inhalt von Bewegungsbeobachtung und Bewegungsvorstellung dar. Es ist sogar davon auszugehen, dass bei diesen beiden Handlungen dieselben neuralen Zentren aktiviert werden und somit eine gute Bewegungsvorstellung die Qualität der Bewegungsbeobachtung maßgeblich beeinflussen kann (Jeannerod, 2001, S. 103f.). Insofern erscheint es lohnend, sich mit dem Konstrukt der Bewegungsvorstellung begrifflich etwas näher zu beschäftigen (siehe Kapitel 2.3.1.1) und Überlegungen anzustellen, wie eine Bewegungsvorstellung entsteht (siehe Kapitel 2.3.1.3).

Insgesamt sind die Forschungsergebnisse, speziell bezogen auf die Sportwissenschaft, jedoch recht diffus. Die grundlagenorientierte Kognitionswissenschaft sowie die Neuropsychologie haben sich in den letzten zwei Jahrzehnten zwar verstärkt dem Bereich der Bewegung und Motorik zugewandt, allzu viele Ergebnisse konnten aber noch nicht auf den sportlichen Kontext übertragen werden. Zumeist geht es in sportwissenschaftlichen Untersuchungen dann um den Einfluss von mentalem Training auf die konkrete Bewegungsausführung eines aktiven Sportlers, weniger um den Wert einer guten Bewegungsvorstellung für Gruppierungen wie Trainer oder Sportlehrkräfte, die Bewegungsausführungen beobachten und entsprechende Rückmeldung darüber weitergeben. In der englischsprachigen Literatur findet seit den 1980er Jahren eine intensivere Auseinandersetzung mit den Vorgängen der Bewegungsvorstellung statt, die jedoch ebenfalls sehr grundlagenorientiert und weniger sportspezifisch ist. Immerhin findet man dort verschiedene Testverfahren über die Qualität von Bewegungsvorstellungen (siehe Kapitel 2.3.1.4) und auch neurowissenschaftliche Untersuchungen, die sich mit dem Entstehen von Bewegungsvorstellungen beschäftigen (siehe Kapitel 2.3.1.2). Die Hauptperspektive ist aber ebenfalls vorwiegend auf den aktiven Sportler und den Zusammenhang zwischen Bewegungsvorstellung und motorischen Lernprozessen oder Leistungsfortschritten gerichtet, weniger auf den Nutzen der Bewegungsvorstellung für beobachtende und rückmeldende Trainer oder Sportlehrer, wie folgende Definition von Daus (1984, S. 41) deutlich macht:

Als Bewegungsvorstellung wird ein bewußtes [sic!], durch ‚chunk‘-Bildung (Bedeutungseinheiten) strukturiertes, generalisiertes begriffliches Konzept (sportlicher) Bewegungen verstanden, das als organisierender Rahmen für ihre jeweils aktuelle umweltbezogene sensomotorische Koordination fungiert.

Trotz dieser einseitigen Forschungsausrichtung kann mit Hale, Seiser, McGuire und Weinrich (2005, S. 131) gefordert werden: „Coaches, you should practice imagery too!“ Insofern sollen im folgenden Kapitel Erkenntnisse über den Prozess der Bewegungsvorstellung aus unterschiedlichen Wissenschaftsbereichen zusammengefasst und entsprechende Folgerungen für die Fragestellung dieser Arbeit gezogen werden, inwiefern die Bewegungsvorstellung ein möglicher Einflussfaktor auf die Beurteilungsleistung sportlicher Bewegungen sein kann.

2.3.1.1 Begriffsbestimmung

Einige Fragen von Annett (1995, S. 1412f.) am Ende seiner Überlegungen über die Bewegungsvorstellung (engl. motor imagery) fasst die Komplexität des Phänomens der Bewegungsvorstellung nachvollziehbar zusammen:

How then can we characterise motor imagery? Is it genuinely motoric, some kind of incomplete action, or is it purely perceptual, a combination of visual and kinesthetic imagery?

Wie korrespondiert also die Bewegungsvorstellung mit den motorischen Programmen, die bei einer konkreten Bewegungsausführung ablaufen, oder welche Rolle spielt die visuelle oder kinästhetische Wahrnehmung beim Vollzug oder der Beobachtung von Bewegungen? Alle diese Fragen zeigen schon, dass man sich der Bewegungsvorstellung von unterschiedlichen Perspektiven nähern kann, auch abhängig davon, ob man auf die Verbesserung der Beobachtungsfähigkeit oder der Ausführungsqualität sportlicher Bewegungen abzielt.

Um zu einem differenzierteren Begriffsverständnis zu gelangen, soll zuerst geklärt werden, was **Vorstellung allgemein** ausmacht, ehe auf den speziellen Charakter der Vorstellung von Bewegungen eingegangen wird. Grundsätzlich definiert Schöpe (1997, S. 9) Vorstellungsvorgänge als „psychische Prozesse, die das Wiederhervorrufen von Sinneseindrücken betreffen, sich aber auch in neuartiger Weise in zusammengesetzter Form wieder erzeugen lassen, ohne daß [sic!] sie unbedingt wirklichkeitsgetreu sein müssen.“ Die Bandbreite von Vorstellungen reicht dabei von konkreten Erinnerungsbildern bis zu reinen Phantasieprodukten. Die Wahrnehmung spielt dabei eine besondere Rolle: Vorstellungen basieren auf wahrnehmungsartigen Erscheinungen und werden auch zu weiten Teilen von den gleichen Gehirnzentren aktiviert, ohne dass aber zwangsläufig externe Wahrnehmungsstimuli vorliegen müssen¹⁶ (Annett, 1995). Außerdem darf eine Vorstellung nicht als Erleben isolierter Sinnesempfindungen oder als Erinnerung von Sinneseindrücken verstanden werden, denn allein die Sinnesregistration stellt noch keine Vorstellung dar. Vielmehr wird die Vorstellung angereichert mit bedeutungshaltigen Gestaltfaktoren, Einstellungen oder Interessen. Die Wirklichkeitstreue einer Vorstellung hängt auch von der Fähigkeit ab, Erlebtes genau rekonstruieren zu können. Dabei sind zum einen nicht alle über die Sinne gemeldeten Daten bewusstseinsfähig, zum anderen werden auch nicht alle bewussten Sinneseindrücke für die Ausbildung einer Vorstellung herangezogen. Im Sport werden die vielen einzelnen registrierten Bewegungsänderungen (wie Beugen oder Strecken von Gelenken, Orts- oder Winkelveränderungen) von einer Vorstellungsstruktur verarbeitet und als interne Abbilder gespeichert, ohne dass zwangsläufig eine **Verbalisation von Wissen** oder deren **biomechanisches Verstehen** miteingeschlossen ist Schöpe (1997, S. 9–13). Gerade diese beiden Fähigkeiten scheinen aber für Trainer und Sportlehrer in ihrer erklärenden und rückmeldenden Tätigkeit besonders wichtig zu sein, weshalb man darauf in der Vermittlung des Bewegungssehens einen besonderen Schwerpunkt legen muss.

¹⁶ Anhand von Untersuchungen von Patienten mit spezifischen selektiven Wahrnehmungsstörungen (z. B. bei Prosopagnosie oder Hemineglect) konnte gezeigt werden, dass auch die Vorstellungsfähigkeit bei solchen Personen gestört ist. Bei geburtsblinden Personen konnte außerdem nachgewiesen werden, dass Prozesse der räumlichen Vorstellung (z. B. bei mentalen Rotationen) trotz fehlender visueller Erfahrungen genauso gut ausgeprägt sind wie bei normal sehenden Personen, was darauf hindeutet, dass die Verarbeitungszentren für Wahrnehmung und Vorstellung zusammenhängen müssen (Munzert (2001b, S.42)).

Betrachtet man nun das Vorstellen von Bewegungen genauer, so taucht übereinstimmend häufig das Kriterium auf, dass es sich um einen **aktiven Prozess** handelt, bei dem eine **geistige Repräsentation von Bewegungen im Arbeitsgedächtnis** internal produziert wird, ohne dass die Bewegungen tatsächlich durch die Muskulatur ausgeführt werden (Decety & Grézes (1999), Jeannerod (2001, S. 103)). Munzert, Lorey und Zentgraf (2009, S. 306) betrachten motor imagery folglich als ein Fenster, welches einen Zugang zu kognitiven motorischen Prozessen und Kontrollmechanismen eröffnen könnte. Wichtig bleibt mit Lotze und Halsband (2006, S. 387) festzuhalten: „The process of imagination is not dependent on the ability to execute a movement but rather on central processing mechanisms.“ Somit kann zwar die Eigenrealisation ein Vorteil für die Ausbildung einer guten Bewegungsvorstellung sein, sie ist aber keine zwingende Voraussetzung.

2.3.1.1.1 Unterteilung der Bewegungsvorstellung

Grundsätzlich kann der Begriff der Bewegungsvorstellung in zweifacher Weise verstanden werden: als Phänomen der Eigenwahrnehmung oder der Fremdwahrnehmung. Wiemann (1989, S. 196) spricht von **eigenmotorischer Vorstellung**, wenn die eigene Körperbewegung gedanklich vollzogen wird, und grenzt davon die Erinnerung an visuelle Eindrücke aus der **Beobachtung einer Fremdbewegung** ab. Jeannerod (2001, S. 103f.) verwendet dafür die englischen Begriffe „imagined action“ und „action observation“. Da es in keinem der beiden Fälle zu einer motorischen Aktion kommt, ist die Bewegungsvorstellung ausschließlich eine Leistung des zentralen Nervensystems und wird somit auch als zentrale Reproduktion (bzw. zentrale Repräsentation) der Bewegung verstanden. Für die Vorstellungsbildung kann ein Beobachter grundsätzlich alle Sinnesmodalitäten (Callow und Hardy (2004, S. 168) sprechen bei Vorstellungsprozessen von imagery modalities) nutzen, was er auch je nach individueller Voraussetzung in unterschiedlicher Weise tut (Munzert, Dültgen & Möllmann, 2000, S. 17), wobei sich nach Annett (1995) im Normalfall die Bewegungsvorstellung über kinästhetische und visuelle Eindrücke ausbildet.

Jeannerod (1994, S. 188f.) wählt für diese Unterscheidung die Begriffe **visual imagery** und **motor imagery**, wobei unter letzterem Begriff eine vorgestellte Bewegung des Körpers (z. B. beim Laufen) oder nur von Körperteilen (z. B. Armbewegung beim Werfen) aus der Eigenperspektive (first-person process) gemeint ist, d. h. der Vorstellende ruft nicht nur seine kinästhetische, sondern auch seine visuell-räumliche Innensicht der Bewegung so ab, als würde er die Bewegung selbst vollziehen. Jeannerod (2001, S. 104) betont, dass Elemente davon bewusstseinsfähig (explicit, conscious) sein können (z. B. die Vorstellung der Bewegung an sich oder das bewusste Beobachten eines anderen Athleten bei seiner Bewegungsausführung), andere wiederum unbewusst (implicit, non-conscious) abgerufen werden (z. B. zukünftige Handlungseinschätzungen oder wahrnehmungsbasierte Entscheidungen). McAvinue und Robertson (2009, S. 114) halten diesbezüglich fest: „explicit and implicit motor imagery are separate abilities“. Die Erkenntnis, dass bei Untersuchungen der visuellen Bewegungsvorstellung kein Zusammenhang festgestellt werden konnte zwischen Selbstauskunftserhebungen und objektivierten Messverfahren, hat dazu geführt, dass zwei Formen der Bewegungsvorstellung unterschieden werden: eine phänomenologische, bewusst erfahrene Vorstellung und eine nicht-phänomenologische Vorstellung, die unbewusst ausgebildet wird. Erstere zeichnet sich dadurch aus, dass die Bewegungsvorstellung mit emotionalen und physiologischen Eindrücken sowie Erinnerungen verknüpft werden kann und offen ist für Interpretationen auf Seiten des Vorstellenden (McAvinue & Robertson, 2009, S. 114).

Während **kinästhetische Vorstellungen** also dabei innere Eindrücke abbilden (z. B. in Form eines individuellen Bewegungsgefühls), die ein Sportler nur während der Eigenrealisation einer Bewegung sammeln kann, beinhaltet die **visuelle Vorstellung** „eine filmartige Sequenz der Bewegung“ (Munzert & Reiser, 2003, S. 220) – Annett (1995, S. 1395) bezeichnet sie auch als „mental video“ – , die der Vorstellende sowohl aus der Außensicht wie auch aus der Innensicht betrachten kann (siehe Abb. 21). Mahoney und Avener (1977, S. 137) unterscheiden als erste diese Perspektiven wie folgt:

In external imagery, a person views himself from the perspective of an external observer (much like in home movies). Internal imagery, on the other hand, requires an approximation of the real-life phenomenology such that the person actually imagines being inside his/her body and experiencing those sensations which might be expected in the actual situation.

Munzert et al. (2009, S. 307) betonen zwar, dass die Außensicht kein Teil des Begriffs motor imagery darstellt, es für Sportler aber kein Problem darstellt, eine solche Außenperspektive zu erzeugen, wenngleich für den Vorstellenden nicht das Gefühl entsteht, als würde er die Bewegung gerade selbst ausführen.

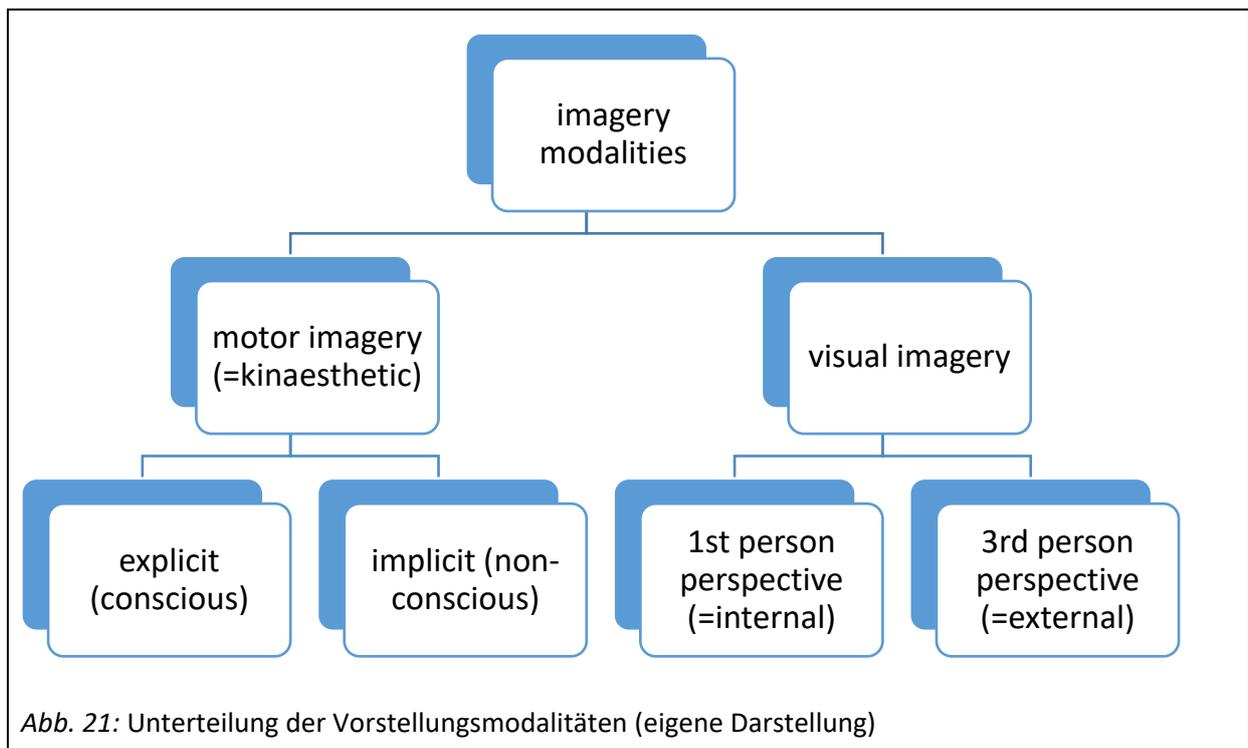
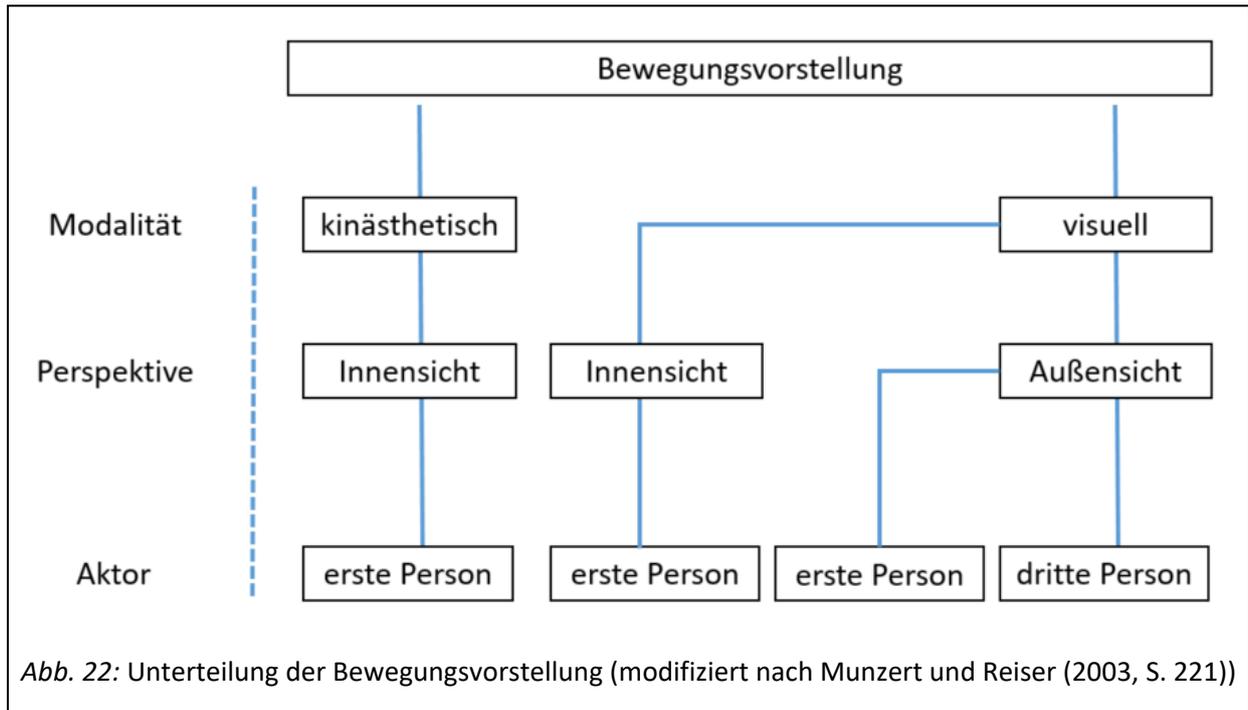


Abb. 21: Unterteilung der Vorstellungsmodalitäten (eigene Darstellung)

Außerdem kann die externe visuelle Bewegungsvorstellung noch derart unterschieden werden, dass der Vorstellende nicht sich selbst von außen, sondern einen Modellathleten bei der Bewegungsausführung betrachtet. Munzert und Reiser (2003, S. 221) unterteilen demnach die Außensicht zusätzlich in eine Erste-Person- und eine Dritte-Person-Perspektive (siehe Abb. 22). Eberspächer (2001, S. 82) spricht dabei von verdecktem Wahrnehmungstraining, wenn „man vor seinem ‚geistigen Auge‘ einen Film über einen Bewegungsablauf, den man selbst ausführt“, ablaufen lässt, sich selbst also aus der Außenperspektive wahrnimmt. Die visuelle Bewegungsvorstellung umfasst in vielen Sportarten außerdem noch komplexe Umweltmerkmale wie Bewegungen von Bällen, Schlägern oder verschiedenen Wurfgeräten sowie Relationen des Körpers zu seiner Umwelt. Es bleibt dabei festzuhalten, dass sich ohne einen Akteur als Träger der Bewegung auch keine Bewegungsvorstellung ausbilden kann (Munzert & Reiser, 2003, S. 220f.).



Insgesamt findet man in der englischsprachigen Literatur sehr viele Bezeichnungen für das Phänomen Bewegungsvorstellung.¹⁷ Um die terminologische Uneinheitlichkeit zu überwinden, bemühte sich Annett (1995) darum, den Begriff **motor imagery** als Vorstellungsfähigkeit zu definieren, die ein gewisses Maß an bewusster Kontrolle von Seiten des Vorstellenden erfordert. Madan und Singhal (2012, S. 223) versuchten, die vielen Begriffe, die auch in verschiedenen Studien noch unterschiedlich verwendet wurden, auf folgende Begriffe zu reduzieren, die sie wie folgt definierten:

- **Visual imagery:** Visuelles Vorstellen von Bewegungen, ohne dass Personen mit Objekten involviert wären, aus der Dritte-Person-Perspektive (external, allozentrisch)
- **Movement imagery:** Vorstellen von einfachen Körperbewegungen von sich selbst aus der Erste-Person-Perspektive (internal) oder von einer anderen Person aus der Dritte-Person-Perspektive (external, allozentrisch)
- **Motor imagery:** Vorstellen der eigenen komplexen Bewegungshandlung von sich selbst, wobei eine dynamische Interaktion mit Objekten erfolgen kann, aus der Erste-Person-Perspektive (internal, egozentrisch)
- **Kinaesthetic imagery:** Bewegungsvorstellung, die ausschließlich auf propriozeptiven Informationen basiert (z. B. Wie fühlt sich die Bewegungsausführung an?), in der Erste-Person-Perspektive, aber weniger bei Ganzkörperbewegungen, sondern bei spezifischen Bewegungen von Körperteilen

¹⁷ McAvinue und Robertson (2008, S.232) nennen als Synonym für imagined movement die Begriffe motor imagery, mental imagery, movement imagery, mental practice, imagery rehearsal, visualisation, kinaesthetic imagery und visuomotor behavioural rehearsal.

2.3.1.1.2 Verhältnis von kinästhetischer und visueller Bewegungsvorstellung

Im Sportbereich kommt es sehr häufig zum Zusammenwirken von visuellen und kinästhetischen Vorstellungsmodalitäten, ohne dass deren Verhältnis zueinander bereits geklärt wäre. Die Innensicht der Bewegungsvorstellung wird beinahe mystifiziert, ohne dass jedoch stärkere kinästhetische Vorstellungen nachgewiesen werden können (Munzert et al., 2000). Laut Hardy und Callow (1999, S. 108) entfallen damit die angeführten Argumente für eine Bevorzugung der Innensicht. Die kinästhetischen Empfindungen bei einer Bewegungsvorstellung können jedenfalls nicht mit dem sog. Carpenter-Effekt erklärt werden, „da bei der Bewegungsvorstellung die im EMG feststellbaren Muskelinnervationen und die Bewegungsreafferenzen sich zwangsläufig von denen bei aktiver Ausführung unterscheiden müssen“ (Munzert, 2001b, S. 44). Somit findet die Vorstellung entsprechender kinästhetischer Empfindungen über einen Gedächtnisabruf und nicht über eine Aktivierung von Muskeln in Abwesenheit der äußeren Kräfte (z. B. der Fliehkräfte beim bogenförmigen Hochsprunganlauf) statt.

Mahoney und Avenier (1977, S. 137) fanden an einer kleinen Probandenzahl aus dem erweiterten US-Olympiateam im Gerätturnen heraus, dass erfolgreiche Athleten häufiger Vorstellungsprozesse aus der internalen Perspektive benutzen als weniger erfolgreiche.¹⁸ Möglicherweise findet dieses Handeln seine theoretische Entsprechung in dem Ergebnis von Hale (1982, S. 384), der herausfand, dass eben diese internale Perspektive stärkere elektromyographische Aktivitäten in der Muskulatur, die an der vorgestellten Bewegung beteiligt ist, hervorruft als eine entsprechende externale Perspektive. Aus diesem Grund vermuteten vor allem die Sportpraktiker, dass Sportler eher auf internale Vorstellungsvorgänge zurückgreifen (z. B. Weinberg (1988, S. 121)), wobei externale Bilder für den Vorstellungsprozess weniger gut geeignet sein sollen als Bilder aus einer internalen Perspektive, da der Sportler sich in der Innensicht seiner eigenen Bewegungsvollzüge besser bewusst wird und mit zusätzlichen Eindrücken anreichert (Hale, 1998, S. 6f.). Ungerleider und Golding (1991, S. 109) jedoch fanden bei erfolgreichen Leichtathleten aus dem erweiterten US-Seoul-Olympiateam eine häufigere Verwendung der externalen Vorstellungsperspektive mit stärkeren physisch nachweisbaren Körperaktivitäten als bei ihren weniger erfolgreichen Kollegen.

Es erscheint nachvollziehbar, dass kinästhetische Vorstellungen mit beiden visuellen Perspektiven in Einklang zu bringen sind (Hardy & Callow, 1999). Callow und Hardy (2004, S. 174) konnten auf der einen Seite einen Zusammenhang zwischen internaler visueller und kinästhetischer Bewegungsvorstellung nachweisen, auf der anderen Seite aber überraschenderweise auch zeigen, dass externale Bewegungsvorstellungen auf kinästhetische Informationen zurückgreifen, vor allem wenn sie in der Dritte-Person-Perspektive ablaufen. Bei Bewegungen, die nach ihrer Form beurteilt werden, hat die kinästhetische Vorstellungsfähigkeit sogar eine stärkere Verbindung zu externen visuellen Vorstellungsinhalten in der Erste-Person-Perspektive als zu internen Vorstellungsinhalten. Hardy (1997, S. 289f.) begründet dies damit, dass z.B. bei Bewegungen wie Turnen oder Klettern eine externale Bewegungsvorstellung den Sportler besser mit zusätzlichen Informationen über die räumlichen Körperbewegungen, den Körperumriss oder Positionen von Gliedmaßen versorgt, die für die Bewegungsausführung wichtig sind, als dies aus einer internalen Perspektive möglich ist.

¹⁸ Callow und Hardy (2004, p.168) kritisieren an dieser Studie, dass die Probanden bezüglich internaler Vorstellungsprozesse lediglich nach ihren Erfahrungen gefragt wurden, wie sich die Vorstellung in der Muskulatur anfühlt, wobei die Phrasen eher auf kinästhetische und motorische Vorstellungsaspekte und weniger auf visuelle bezogen waren.

Jeannerod (2001, S. 104) betont, dass durch die Bewegungsbeobachtung aus der Außensicht die technischen Bewegungsmerkmale besser simuliert werden können, wodurch gleichzeitig auch das Reproduzieren dieser Bewegungsmerkmale erlernt wird. Die Entscheidung zwischen Innen- und Außensicht muss demnach abhängig von der Art der Aufgabe getroffen werden, eine Bevorzugung der Innensicht, wie sie häufig aus sportpraktischer Sicht abgeleitet wird, ist empirisch jedenfalls nicht begründbar (Munzert, 2001a, S. 47). McAvinue und Robertson (2009, S. 114f.) bezeichnen die kinästhetische Vorstellung als dynamischen Prozess, der von vielen Faktoren beeinflusst ist, während der rein visuelle Vorgang einen eher statischen Prozess darstellt.

Wenn der Fokus auf der Beobachtungsfähigkeit von Bewegungen liegt und Trainer oder Sportlehrkräfte die Zielgruppe von Lernprozessen des Bewegungssehens sind, so hat man mit der Schwierigkeit zurecht zu kommen, dass das Eigenrealisationsniveau jedes Beobachters unterschiedlich sein wird, je nach Vorerfahrung in der entsprechenden Sportart. Somit sind die kinästhetischen Erfahrungen, welche eine Bewegungsvorstellung maßgeblich mitprägen können, in höchst unterschiedlichem Maße vorhanden und mitunter schwer nachzuholen, wenn man nicht schon als junger Sportler in der entsprechenden Sportart aktiv war. Daher soll im Rahmen dieser Arbeit vor allem die visuelle Bewegungsvorstellung interessieren, die ein Beobachter in der Außenperspektive durch die Bewegungswahrnehmung einer dritten Person gewinnt. Dazu gehört nicht nur das Wissen um die perfekte Bewegung, sondern auch um altersgemäße Abweichungen von diesem Idealbild.

2.3.1.2 Neurophysiologische Grundlagen

Durch die Entwicklung neuer bildgebender Verfahren in den 1990er Jahren haben auch die neurophysiologischen Erkenntnisse über die Bewegungsvorstellung zugenommen. Munzert et al. (2009, S. 308f.) führen 43 Studien für den Zeitraum von 1995 bis 2008 an, die mit Hilfe neurophysiologischer Methoden das Phänomen Bewegungsvorstellung genauer untersucht haben. Vor allem befassten sich viele Autoren mit der Hypothese, dass die Bewegungsvorstellung auf denselben Strukturen basiert wie die Bewegungsplanung bzw. -vorbereitung oder als Bewegungsausführung mit gehemmtem Endglied zu verstehen ist. Parsons (1994, S. 709) stellt allgemein fest: „sensorimotor structures support mental simulations of actions“. Genau in diese Richtung argumentiert auch Jeannerod (2001, S. 103), wenn er im Rahmen der *Simulation theory* beschreibt, dass vorgestellte Bewegungen auf den gleichen neuronalen Prozessen beruhen wie reale Bewegungen, nur dass sie nicht ausgeführt werden:

The simulation theory [...] postulates that covert actions are in fact actions, except for the fact that they are not executed. The theory therefore predicts a similarity, in neural terms, between the state where an action is simulated and the state of execution of that action.

Für diese neuronale Simulation der Handlung, welche große Übereinstimmung mit der Handlungsausführung hat, prägt Jeannerod (2001, S. 103) den Begriff **S-state**. Jedoch können auch Unterschiede festgestellt werden zwischen der Bewegungsvorstellung und der Bewegungsausführung, weil motor imagery nur die verborgene Stufe der Handlungsausführung ist, während die motorische Ausführung zusätzlicher Prozesse auf der beobachtbaren Verhaltensebene bedarf (Munzert et al., 2009, S. 310). Für visuelle Vorstellungsinhalte kann angenommen werden, dass sie auf denselben neuronalen Strukturen beruhen wie reale Bilder, die über das optische System wahrgenommen werden. Die visuelle Repräsentation ist dabei hierarchisch organisiert, d. h. es gibt unterscheidbare Prozesse, je nachdem, ob die Bilder auf der Basis von Wahrnehmung oder Erinnerung generiert werden (Jeannerod, 1994, S. 188f.).

Decety (1996, S. 91) erklärt die gehemmte Bewegungskontrolle damit, dass die Aktivierung des Kleinhirns während der Bewegungsvorstellung inhibitorische Wirkung hat, indem es efferente Impulse, die durch den Vorstellungsprozess ausgelöst werden, an das Rückenmark und die Muskulatur verhindert – wenngleich nicht verschwiegen werden darf, dass es selbst bei normalen Bewegungskontrollen auch solche inhibitorischen Prozesse gibt, welche beispielsweise irrelevante Programmteile hemmen. Kasess et al. (2008) fanden heraus, dass einige neuronale Strukturen des supplementär-motorischen Areals (SMA), welche nur teilweise mit den an der Bewegungskontrolle beteiligten SMA-Strukturen übereinstimmen, die Aktivität von M1 hemmen und somit die motorische Ausführung einer Bewegungsvorstellung verhindern. Es kann festgehalten werden, dass Bewegungsvorstellung und motorische Aktion sowohl den prämotorischen Cortex (PMC) als auch das SMA aktivieren, jedoch keine vollständige Übereinstimmung der neuronalen Aktivitäten vorliegt (Munzert et al., 2009, S. 312).

Betrachtet man die neurophysiologischen Vorgänge von der Bewegungsplanung bis hin zur konkreten Ausführung, so spielen dabei cortikale wie subcortikale Gehirnbereiche eine wichtige Rolle, wobei folgende Strukturen der Großhirnrinde (engl. cerebral cortex) unterschieden werden: der primäre motorische Cortex (M1, Area 4 nach Brodmann), verschiedene prämotorischen Felder (Area 6 nach Brodmann) wie das supplementär-motorische Areal (SMA, medialer Teil der Area 6), das präsupplementär-motorische Areal (pre-SMA, unmittelbar rostral zur SMA angrenzend) und der prämotorische Cortex (PMC, lateraler Teil der Area 6) mit seinem dorsalen und ventralen Teil (siehe Abb. 23). Die prämotorischen Felder haben eine große Bedeutung für die Bewegungsplanung, die Entwicklung von Bewegungsstrategien, die Initiierung von Bewegungsabfolgen und speziell das supplementär-motorische Areal für beidhändige Aktionen, während der primäre motorische Cortex für das Einleiten und Auslösen einfacher Bewegungen verantwortlich ist (Munzert, 2001b, S. 49). Bereits Roland, Larsen, Lassen, und Skinhoj (1980) fanden heraus, dass auch bei Bewegungsvorstellungen die prämotorischen Felder aktiviert werden. Einige Jahre später konnte auch eine Aktivität des primären motorischen Cortex bei Bewegungsvorstellungen nachgewiesen werden (Schnitzler, Salenius, Salmelin, Jousmäki & Hari, 1997).¹⁹ In einer experimentellen Studie von Lotze et al. (1999, S. 494) konnte beim Vergleich von ausgeführten Fingerbewegungen mit der reinen Vorstellung der Bewegungskontrolle eine gleich große Aktivierung der prämotorischen Felder festgestellt werden, wohingegen die Aktivierung des primären motorischen Cortex bei der Vorstellung geringer ausfiel. Beachtlich war auch die Erkenntnis, dass die Lateralisierung der dominanten Hand, welche bei der Ausführung stärker aktiviert wurde, bei der Bewegungsvorstellung nicht mehr nachweisbar war.

¹⁹ Auch hier ist ein Blick auf die medizinische Forschung wieder interessant. Erland et al. (1996) fanden bei armamputierten Personen heraus, dass die Vorstellung von Fingerbewegungen mit der amputierten Hand zur Aktivierung des kontralateralen motorischen Cortex führt, allerdings nur mit etwa 30-prozentigem Aktivierungsniveau im Vergleich zur Ausführung mit der intakten Hand.

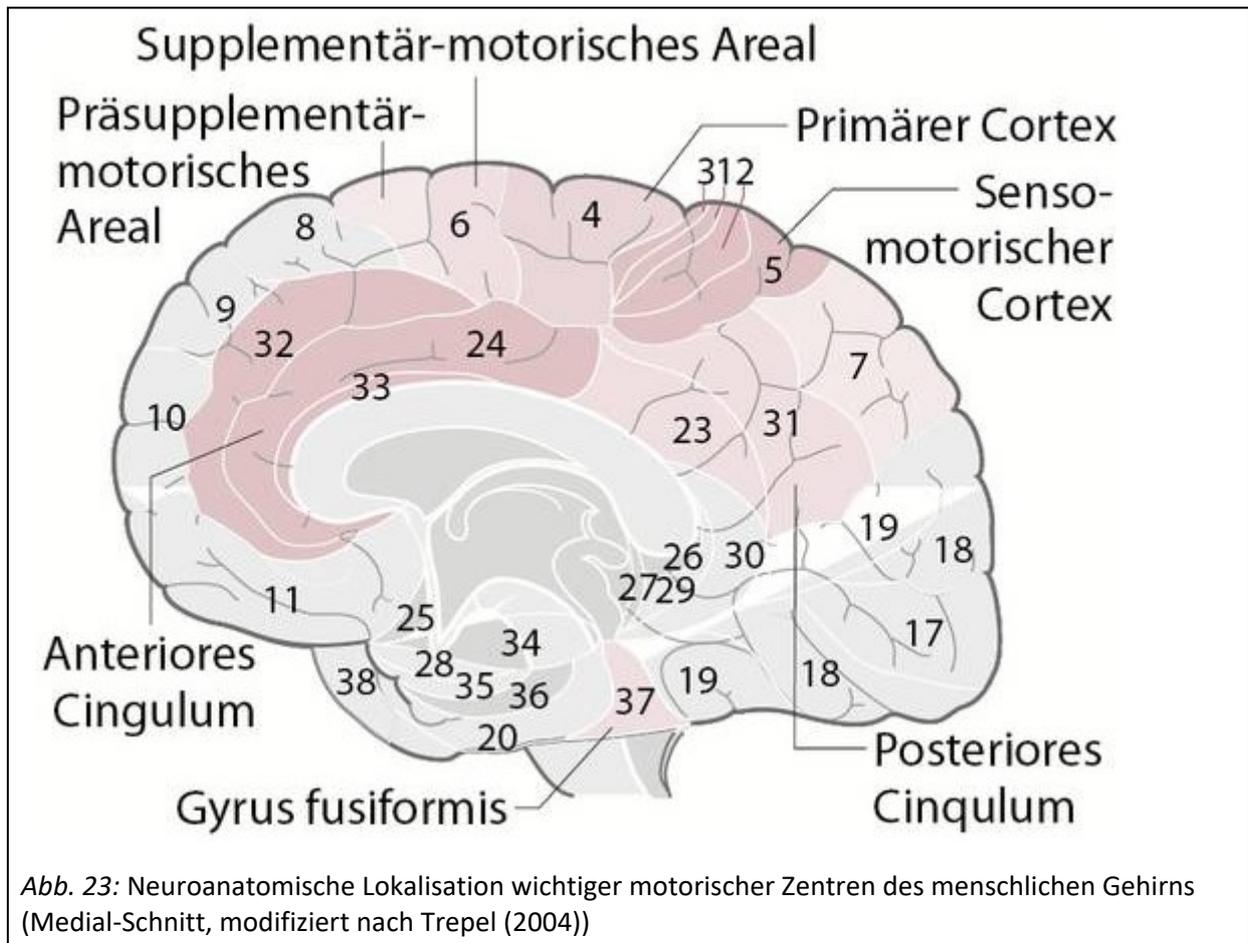


Abb. 23: Neuroanatomische Lokalisation wichtiger motorischer Zentren des menschlichen Gehirns (Medial-Schnitt, modifiziert nach Trepel (2004))

Neben den neuronalen Prozessen kommt es bei der Bewegungsvorstellung auch im peripher-physiologischen Bereich zu auffälligen Übereinstimmungen gegenüber der Bewegungsvorbereitung. Bei der Bewegungsvorstellung steigen die Herz- und Atemfrequenz ebenso wie bei der Vorbereitung auf aktive Übungsausführung, fallen jedoch um ein Viertel geringer aus. Hier muss es sich um unbewusste Prozesse handeln, da das autonome Nervensystem nicht willkürlich gesteuert werden kann. Decety (1996) erklärt sich diese Erscheinung damit, dass der Körper – verleitet durch cerebrale Vorgänge – annimmt, dass Bewegungen auch wirklich ausgeführt werden. Somit zeigt sich beim Vorstellen von Bewegungen auch der Stellenwert zentraler cerebraler (und zugleich unbewusster) gegenüber peripherer Anteile bei der Beeinflussung der Herzfrequenzrate (Munzert, 2001b, S. 50).

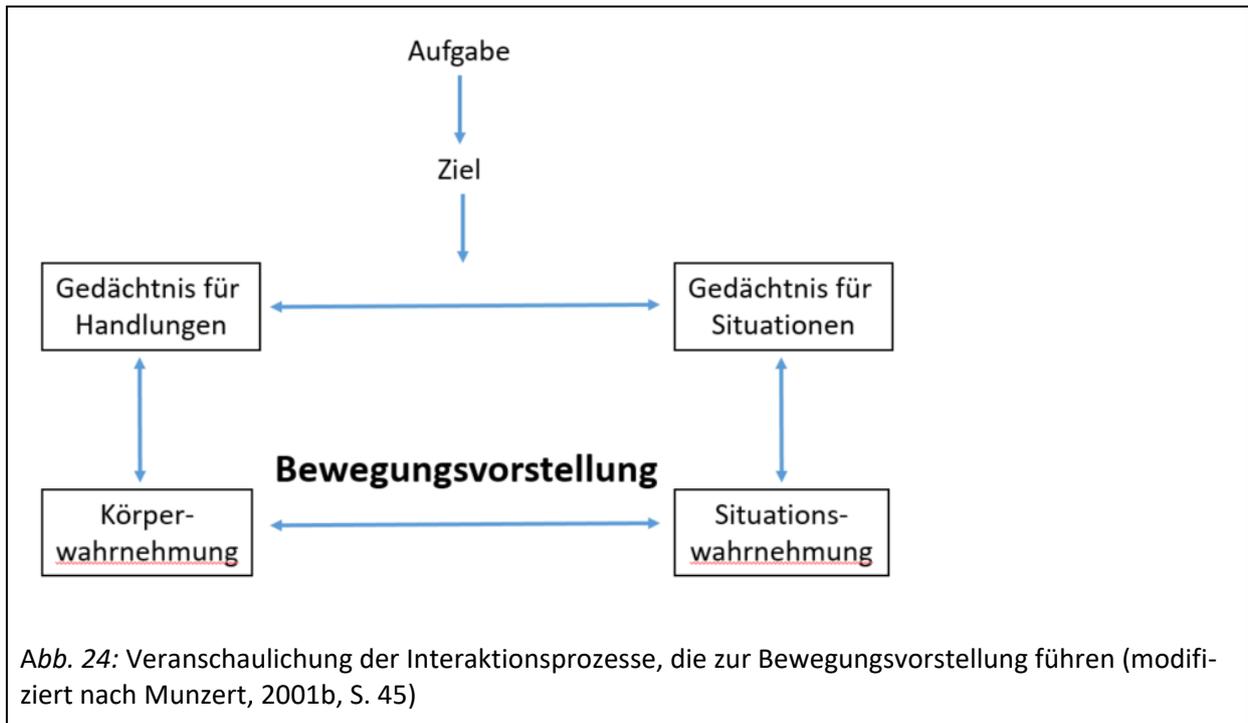
Aus angeführten Studien kann demnach schlussgefolgert werden, dass eine vorgestellte Bewegung die gleiche zeitliche Charakteristik aufweist wie die tatsächlich ausgeführte Bewegung und zum Teil auch dieselben neuronalen wie peripher-physiologischen Bereiche aktiviert werden, jedoch bei der Bewegungsvorstellung mitunter in geringerer Stärke oder mit unterschiedlichen Schwerpunkten (z. B. innerhalb des Kleinhirns). Trotzdem sollten die Prozesse von Bewegungsausführung und Bewegungsvorstellung nicht unhinterfragt gleichgesetzt werden. Während Decety, Jeannerod und Prablanc (1989) in ihrer Grundlagenforschung den engen Zusammenhang zwischen ausgeführter und vorgestellter Bewegungsdauer beim Durchlaufen verschieden langer Laufstrecken als Beleg für eine funktionale Äquivalenz von Vorstellung und Ausführung sehen, betont Munzert (2001b, S. 54) den selektiven Einfluss verschiedener Person-, Aufgaben- und Umweltmerkmale auf den Vorstellungsprozess und hebt darauf ab, die Vorstellung als einen eigenständigen kognitiven Prozess zu betrachten, der dann auch eigenen Gesetzmäßigkeiten und Verläufen folgt.

2.3.1.3 Entstehen von Bewegungsvorstellungen

Die bloße Aufforderung, sich eine sportliche Bewegung vorzustellen, wird so manchen Sportler wie auch Trainer überfordern. Eine entsprechende Vorstellungsfähigkeit ist somit notwendig, um Bewegungen und Handlungen lebhaft und intensiv vor dem inneren Auge ablaufen lassen zu können. Voraussetzung dafür ist wiederum ein **Wissen über die Bewegung** an sich, für welche der finale Bezug der Bewegung eine zentrale Rolle spielt. Gemeint sind damit die Sinn-, Zweck- und Zielsetzungen, die entweder für einzelne Bewegungen oder ein größeres sportliches Handlungsgefüge präsent sind und die Frage klären, was man mit einer bestimmten Bewegung erreichen oder wie man ein gewünschtes Ergebnis herbeiführen kann. Somit ist ein Wissen über die zu bewältigende Bewegungsaufgabe ebenso wichtig wie das Wissen über situative Bedingungen des Bewegungsvollzugs. Als dritter Faktor kommt noch das Wissen über die eigene Person (den Akteur selbst) hinzu, wenn es darum geht, beispielsweise Fehler bei der Bewegungsausführung zu erkennen und zu Umweltbedingungen in Beziehung zu setzen. Das Wissen über die Funktionsweise der Bewegung sowie deren einzelne Bestandteile, deren zeitliche Segmentierung oder das spezifische Bewegungsempfindungen ist in den Motorik-Repräsentationen implizit enthalten (Munzert, 1992, S. 350f.). Zugleich stellt sich auch die Frage, wie realitätsgetreu die Vorstellung (z. B. im Rahmen des mentalen Trainings) sein soll. Der oft gehörte Grundsatz *Je mehr (detailgetreuer), desto besser* muss jedoch angesichts der Forschungsergebnisse über Bewegungsvorstellung infrage gestellt werden und in Abhängigkeit von der jeweiligen Zielsetzung der Bewegungsvorstellung modifiziert werden (Munzert, 2001b, S. 54).

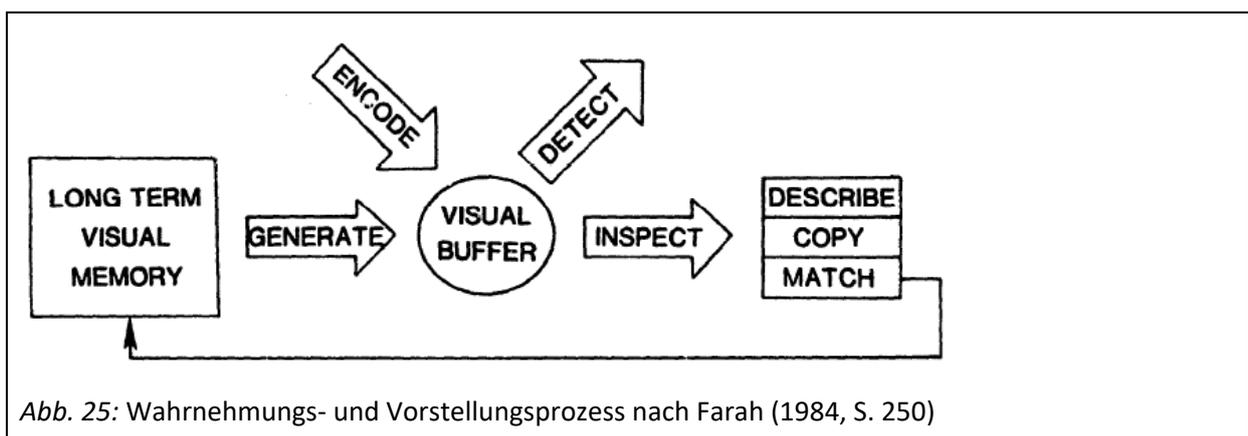
Die Frage, wie sich eine Bewegungsvorstellung ausbildet, wird in der sportwissenschaftlichen Literatur häufig im Kontext des mentalen Trainings im Sinne einer planmäßig wiederholten, bewussten Vorstellung von Bewegungsabläufen thematisiert (Ulich, 1967). Munzert (2001b, S. 45) schlägt vor, diese Definition um die Zielstellungen Lernen, Optimieren und Stabilisieren der vorgestellten Bewegungsabläufe zu erweitern, denn davon wird die Qualität des Vorstellungsprozesses maßgeblich beeinflusst. Um eine Bewegungsvorstellung zu entwickeln oder zu verbessern, können entsprechende Verfahren zur Anwendung kommen, mit deren Hilfe motorische Abläufe bewusst gemacht werden. Dabei stellt sich grundsätzlich die Frage, ob Bewegungsvorstellungen ein Derivat motorischer Programme sind und somit auf Basis wiederauflebender sensorischer Eindrücke entstehen oder über bedeutungshaltige, episodische Sequenzen rekonstruiert werden. Munzert (2001a, S. 57) spricht sich daher gegen die Bewegungsvorstellung als Folge der Aktivierung und Hemmung von Motorikprogrammen aus, weil Bewegungsvorstellungen (z. B. mittels kinästhetischer Empfindungen) generell reichhaltiger sind als die Inhalte des Motorikprogramms.

Für die Vorstellungsbildung muss keine eigene Instanz angenommen werden, sondern es werden wahrnehmungsgebundene Gedächtnisinhalte abgerufen, welche im Rahmen der neuronalen Wahrnehmungsstrukturen die modalitätsspezifischen Vorstellungen hervorrufen. Die Bewegungsvorstellung entsteht demnach aus der Interaktion von Wahrnehmungs- und Gedächtnisprozessen und wird konzipiert als Produkt der Interaktion von Inhalten des Situations- und Handlungsgedächtnisses, welche untereinander stark verbunden sind und wiederum mit den entsprechenden Prozessen der Situations- und Körperwahrnehmung in Verbindung stehen (siehe Abb. 24). Vorstellung und Handlungsvorbereitung sind somit funktional äquivalent, denn auch die Handlungsvorbereitung beruht auf der Interaktion der eben dargestellten Funktionen. Die Auswahl und Initiierung der Aktionen schafft dann den Übergang vom Vorstellungsprozess zum konkreten Handlungsvollzug. Erfolgt diese Stufe nicht, verbleibt es bei der Vorstellungsbildung im eigentlichen Sinne (Munzert, 2001b, S. 44f.).



2.3.1.3.1 Das Information-Processing Model of Imagery nach Farah

Farah (1984, S. 250) schlägt ein vereinfachtes Grundmodell des visuellen Vorstellungsprozesses (*Information-Processing Model of Imagery*) vor, bei dem Inhalte des visuellen Langzeitgedächtnisses in das Arbeitsgedächtnis transformiert werden und dort wahrnehmungsartige Erscheinungen der Vorstellung hervorrufen bzw. generieren (siehe Abb. 25). Bevor die Vorstellungsinhalte im Arbeitsgedächtnis wieder verblassen, kann die Vorstellung auch verändert bzw. transformiert werden, indem beispielsweise erwünschte Konstellationen der Bewegungsrealisation vorweggenommen werden, die dem Handeln dann als Zielvorstellungen dienen. Wenn man sich seine Vorstellung bewusst macht und damit auf direktem oder indirektem Weg anderen mitteilt, spricht Farah von Inspektion. Angewendet auf sportliche Bewegungen kann dieser Vorgang auf dreifache Weise erfolgen: 1. Spezifische Inhalte der Bewegungsvorstellung werden beschrieben, 2. bestimmte Vorstellungsmerkmale werden physisch kopiert (z. B. wenn ein vorgestellter Gelenkwinkel mittels einer Gliederpuppe nachgebildet wird) und 3. aktuelle Vorstellungen werden mit einem länger zurückliegenden Sachverhalt verglichen (z. B. wenn ein aktuell ausgeführter Bewegungsinhalt von vorherigen Versuchen oder einem Modell abweicht) (Munzert, 2001b, S. 42f.).



In diesem semantisch geprägten, auf neuropsychologischen Befunden basierenden Funktionsschema geht Farah (1984) also von der Tatsache aus, dass zur Vorstellungsbildung zuerst „Informationen aus dem Langzeitgedächtnis in ein Arbeitsgedächtnis überführt werden und dabei die wahrnehmungsartigen Erfahrungen der Vorstellung hervorrufen“ (Munzert, 2001a, S. 54f.). Obwohl noch nicht geklärt ist, wie man an die richtige Information im Gedächtnis gelangt, so funktioniert dieser Prozess (vermutlich über verbale Selbstinstruktionen) recht zuverlässig, vor allem, wenn die Aufmerksamkeit (beispielsweise durch Schließen der Augen) nach innen gerichtet wird. Demnach tragen drei unterschiedliche Teilsysteme, die auch individuell verbesserbar sind, zu einem erfolgreichen Vorstellungsprozess bei:

- **Beschreiben:** Übersetzung des Bewegungsablaufs in einen verbalen Code
- **Kopieren:** Umsetzen der Vorstellung in eine graphische, visuell nachvollziehbare Form (z. B. Darstellung von Körperpositionen mittels einer Gliederpuppe)
- **Vergleichen:** mit anderen Gedächtnisinhalten (unter Rückbezug zum Langzeitgedächtnis)

Diese drei Prozesse können als Grundprinzipien dienen, um die Bewegungsvorstellung von Trainern und Lehrkräften methodisch-didaktisch zu verbessern.

2.3.1.3.2 Der sprachlich-symbolische Ansatz

Mayer und Hermann (2011, S. 26–28) nehmen das **Stufenmodell des Mentalen Trainings** (nach Eberspächer, 2001, S. 86f.) zum Ausgangspunkt für seinen **sprachlich-symbolischen Ansatz** der Bewegungsvorstellung. Grundlegend werden die „Vorstellungsinhalte verbalisiert, meist schriftlich fixiert, auf wesentliche Punkte reduziert und dann dem Bewegungsrhythmus angepasst“ (Mayer & Hermann, 2011, S. 26). Dieser Prozess erfolgt in vier Stufen:

1. Detaillierte Beschreibung der Bewegung
2. Hervorhebung der Knotenpunkte
3. Symbolische Markierung und Rhythmisierung der Knotenpunkte
4. Mentales Training der symbolisch markierten und rhythmisierten Knotenpunkte

Als erste Stufe der Generierung einer Bewegungsvorstellung ist es demnach wichtig, eine detaillierte Beschreibung der Bewegung im Gedächtnis verfügbar zu haben, die in mündlicher oder schriftlicher Form reproduziert werden kann und den Bewegungsablauf in möglichst vielen Sinnesmodalitäten nachvollziehbar darstellt. Dabei dürfen und sollen sogar subjektive Erlebensprozesse Eingang in die Bewegungsvorstellung finden. Auf dieser Grundlage der Bewegungsbeschreibung werden dann zentrale Knotenpunkte²⁰ des Bewegungsablaufs, die unbedingt durchlaufen werden müssen, systematisiert und so die Bewegungsstruktur deutlich gemacht. Wo diese entscheidenden Stellen einer Bewegung liegen, hängt dabei auch vom Könnensstand der jeweiligen Zielgruppe ab, die beobachtet werden soll. Diese Knotenpunkte werden anschließend symbolisch markiert, d. h. sie werden in individuelle Kurzformeln²¹ umbenannt, wodurch eine Verdichtung der Information (sog. chunking) erfolgt. Die Be-

²⁰ Beim Tennisaufschlag formuliert Eberspächer (2001, S.87) beispielsweise folgende Knotenpunkte: 1. Ballhochwerfen, 2. Bogenspannung, 3. Schlägerrückführung, 4. Streckung und 5. Abklappen des Schlägers. Wenn ein vorheriger Knotenpunkt nicht korrekt durchlaufen wird, hat dies auch negative Auswirkungen auf die folgenden Knotenpunkte.

²¹ Für einen Diskuswerfer könnte man die Knotenpunkte wie folgt symbolisch markieren: „Laaang (Eindrehen) – Sta (Landung zur Wurfauslage) – Pap (Abwurf)“ (Eberspächer (2001, S.87)).

wegungsvorstellung kann somit an den dynamischen und zeitlichen Verlauf der Realbewegung angepasst werden. Wenn der Trainer diese Bewegungsvorstellung an seine betreuten Sportler vermittelt, so können auf dieser Basis Korrekturen exakter formuliert werden.

2.3.1.3.3 Der räumlich-bildhafte Ansatz

Beim **räumlich-bildhaften Ansatz** wird die Bewegungsvorstellung mithilfe visueller Medien wie Videoaufzeichnungen entwickelt. Durch das Beobachten einer anderen Person kann selbst bei bisher noch unbekanntem Techniken eine räumlich-zeitliche Vorstellung der richtigen Bewegungsausführung ausgebildet werden. Durch erweiterte Analysemöglichkeiten wie der Übereinanderlegung von Bewegungsabläufen können kleine Bewegungsvariationen erkannt und analysiert werden (Mayer & Hermann, 2011, S. 28f.).

Häufig wird in der praktischen Anwendung eine Kombination [...] genutzt. So werden im Sport beispielsweise Bewegungsbeschreibungen verfasst, und man versucht, diese durch intensives Videostudium zu differenzieren. Häufige Nachfragen des Trainers nach gelungenen Handlungssequenzen und den Empfindungen des Athleten während des Bewegungsablaufs sollen schließlich dazu beitragen, dass die Bewegungsvorstellung der aktuellen Leistungsentwicklung stetig angepasst wird. (Mayer & Hermann, 2011, S. 30)

Bezieht man sich auf die Ergebnisse der mentalen Chronometrie (siehe Kapitel 2.3.1.4), so bleibt festzuhalten, dass Vorstellungen gegenüber der realen Ausführung den Vorteil haben, dass sie nach gewisser Übung in unterschiedlicher Geschwindigkeit realisiert werden können. Es sollte also grundsätzlich das Ziel sein, eine Bewegung in der Vorstellung sowohl in Echtzeit als auch in verlangsamter Geschwindigkeit ablaufen lassen zu können. Vor allem für das Einprägen oder Verändern von Einzelheiten des Bewegungsablaufs kann dies hilfreich sein. Dauert eine Bewegungsvorstellung hingegen aufgrund der Komplexität der Bewegung wesentlich länger als die real durchgeführte, so können Vorstellungsinhalte durch die Bildung von (weiteren) Knotenpunkten zusammengefasst und somit auch die Vorstellung beschleunigt werden. Wichtig bleibt dabei, dass Veränderungen der mentalen Bewegungsdauer gleichmäßig alle Bewegungsteile umfassen und das zeitliche Verhältnis der Gesamtbewegung erhalten bleibt (Konzept des relativen Timings). Es gilt aber auch zu bedenken, dass ein solcher Zeitlupenablauf schwer in die kinästhetische Vorstellung zu integrieren ist und auch die Übertragbarkeit auf die motorische Programmierung in Frage steht (Munzert, 2001b, S. 54).

2.3.1.3.4 Der PETTLEP-Ansatz nach Holmes und Collins

Das Akronym PETTLEP steht für die Begriffe Physical Environment Task Timing Learning Emotion Perspective. Dieser Ansatz betont die motorischen und kinästhetischen Elemente der Bewegungsvorstellung. Die physikalische Komponente beschreibt, dass sich die Vorstellung eng an die Sinneseindrücke während der physikalischen Bewegungsvorbereitung und -ausführung anlehnen sollte. Die Umgebung, in welche die Vorstellung eingebettet werden soll (environment), sollte möglichst realistisch zu den Trainings- und Wettkampfstätten passen, welche der Athlet kennt oder worauf er sich vorzubereiten versucht. Der Vorstellungsprozess sollte dabei auf die entsprechende Zielgruppe und Aufgabe abgestimmt sein (task), was beispielsweise das Fertigniveau des Athleten, seinen Lernstand oder die Komplexität der Bewegungsaufgabe umfasst. Sportliche Bewegungen, bei denen die äußere Form eine

wichtige Rolle spielt (z. B. beim Turnen), sollten beispielsweise eher in einer externen Perspektive visualisiert werden, weil die Athleten dann besser die von außen beobachtbaren Bewegungsmerkmale wahrnehmen können, die später dann auch von Kampfrichtern beurteilt werden. Bezüglich der Ausführungsdauer der vorgestellten Bewegung (timing) sollte die Abweichung von der realen Ausführungszeit möglichst gering sein. Ebenso sollte beim Vorstellungstraining berücksichtigt werden, dass sich die Vorstellungsmethode an das sich stetig ändernde Können der Adressaten anpassen muss (learning). Dies kann beispielsweise der Fall sein, wenn die externe Perspektive, die vor allem beim Erlernen neuer Bewegungen vorteilhaft ist, zugunsten der internalen Perspektive aufgegeben wird, welche eher auf kinästhetische Bewegungsaspekte fokussiert ist (perspective). Je emotionsreicher die Bewegungsvorstellung dabei gestaltet werden kann, desto wirkungsvoller wird sie in Wettkampfsituationen eingesetzt werden können (emotion) (Holmes & Collins, 2001).

2.3.1.4 Operationalisierung der Bewegungsvorstellung

Grundsätzlich lässt sich mit McAvinue und Robertson (2009, S. 115f.) festhalten, dass die Vorstellungsfähigkeit keine menschliche Basisfähigkeit ist, die jedermann in gleichem Maße verfügbar hat, sondern individuell messbare Unterschiede vorliegen, die darauf hindeuten, dass die Vorstellungsfähigkeit erlernbar ist. Folgende Äußerung von Jeannerod und Decety (1995, S. 727) scheint diesbezüglich zumindest in Frage gestellt werden zu dürfen: „Motor imagery is a cognitive state that can be experienced by virtually everyone with minimal training.“ Die Schwierigkeit einer Operationalisierung liegt darin, dass die Vorstellungsprozesse an sich nicht beobachtet, sondern nur indirekt über das motorische Ergebnis oder neuronal messbare Anpassungserscheinungen beurteilt werden können (Munzert et al., 2009, S. 309). Wohl aber halten Daus, Blischke, Olivier und Marschall (1989, S.48) fest, dass sich Bewegungsvorstellungen sowohl sprachlich fassen als auch in Vorstellungsbildern aktualisieren lassen und auf diese Weise auch messtechnisch erfasst werden können.

Die Möglichkeiten zur Erfassung der Bewegungsvorstellung sind jedoch noch vielfältiger und werden von Munzert (2001b, S. 51f.) in drei Ebenen unterteilt: die phänomenale Ebene, die Verhaltensebene und die (neuro-)physiologische Ebene. Vorweg bleibt festzuhalten, dass die verschiedenen Ebenen kaum aufeinander zu beziehen sind, sondern auf sehr unterschiedliche Art und Weise die Bewegungsvorstellung zu operationalisieren versuchen. Die **(neuro-)physiologische Ebene** setzt vor allem bildgebende Verfahren ein, um Aktivitäten spezifischer Gehirnregionen bei Vorstellungsprozessen von Bewegungen festzustellen. Ein direkter Zugang zu den Vorstellungsprozessen selbst ist jedoch auch nicht möglich, allenfalls die Vermutung eines korrelativen Zusammenhangs. Bei der **Verhaltensebene** geht es um die Erfassung äußerlich sichtbaren Verhaltens, das als Folge einer Bewegungsvorstellung eine besondere beobachtbare Ausprägung erhält. Diese Methode wird v.a. bei Wirksamkeitsuntersuchungen des mentalen Trainings eingesetzt, ohne dass jedoch die Vorstellungsprozesse an sich thematisiert werden, sondern lediglich die Wirkung von Vorstellungsinstruktionen auf den motorischen Lernerfolg untersucht werden. Ein Beispiel dafür stellt Blischke (1986) dar, der die Auswirkungen eines Bewegungsvorstellungstrainings zum Erlernen der großen Körperwelle auf die motorische Ausübung dieser Fähigkeit untersucht hat (siehe Kapitel 2.3.1.4.5). Grundlage der **phänomenalen Ebene** ist die subjektive Erfahrung einer individuell spezifischen, wahrnehmungsartigen Erscheinung, die auf verschiedene Weise anderen Personen mitgeteilt werden kann. Methodisch kommen hier Interviewverfahren (wie z. B. bei Munzert et al., 2000) oder Selbsteinschätzungsskalen zur Effektivität der eigenen Bewegungsvorstellung (wie z. B. beim **Movement Imagery Questionnaire** (MIQ) von Hall und Pongrac (1983)) zum Einsatz (siehe Kapitel 2.3.1.4.1).

2.3.1.4.1 Selbsteinschätzungsfragebögen

Beim MIQ-R, einer revidierten und vereinfachten Form des Originaltests von Hall und Martin (1997), müssen die Versuchspersonen zunächst eine vorgegebene Bewegung aktiv ausführen, in die Ausgangsposition zurückkehren und dann die Bewegung gedanklich nachvollziehen. Die Fähigkeit, sich die entsprechenden Bewegungen der oberen bzw. unteren Extremitäten oder Ganzkörperbewegungen vorstellen zu können, wird dann auf einer Likert-Skala von 1 (sehr leicht vorzustellen/fühlen) bis 7 (sehr schwer vorzustellen/fühlen) von den Probanden eingeschätzt. Dabei werden die kinästhetische und die visuelle Vorstellungsfähigkeit mit jeweils vier Aufgabenstellungen getestet. Der MIQ ist empirisch gut abgesichert und erreichte bei Untersuchungen von Hall, Pongrac und Buckolz (1985) sowie Atienza, Balaguer und Garcia-Merita (1994) einen Reliabilitätskoeffizienten von $r = .87$ und $r = .89$ für die visuellen und $r = .91$ und $r = .88$ für die kinästhetischen Einschätzungen sowie einen Test-Retest-Koeffizienten von $r = .83$ in beiden Studien. Problematisch an diesem Testverfahren ist allerdings, dass die Daten allein von der Objektivität und Selbstwahrnehmung der Probanden abhängen und außerdem nur bewusstseinsfähige Prozesse erfasst werden (Moreau, Clerc, Many-Dannay & Guerrien, 2010, S. 95).

Ein ebenfalls in der Vergangenheit sehr häufig eingesetzter Test ist der **Vividness of Movement Imagery Questionnaire (VMIQ)**, den die Autoren Isaac, Marks und Russell (1986) mit einer nachgewiesenen Test-Retest-Reliabilität von $r = .76$ entwickelt haben, um die visuelle und kinästhetische Vorstellungsfähigkeit verschiedener sportmotorischer Aufgaben zu messen. 24 Items werden dabei in sechs Gruppen unterteilt: Grundbewegungen (Items 1–4), Grundbewegungen mit erhöhter Bewegungsgenauigkeit (Items 5–8), kontrollierbare Bewegungen mit gelegentlichen Unplanbarkeiten (Items 9–12), Bewegungen mit zu kontrollierenden Objekten (Items 13–16), Bewegungen mit Gleichgewichts- und Erholungsanforderungen (Items 17–20) und Bewegungen mit verschiedenen langen Flugphasen (Items 21–24). Verschiedene Folgeuntersuchungen haben herausgefunden, dass ein positiver Zusammenhang besteht zwischen dem VMIQ-Score und der motorischen Leistungsfähigkeit von Sportlern (z. B. zwischen Profis und Amateuren) (z. B. Hardy & Callow, 1999). Dean (2003, S. 247f.) hingegen kritisiert, dass ein Phänomen wie Klarheit (engl. vividness) nicht das alleinige Maß für die Vorstellungsfähigkeit sein kann, weil es viele andere Variablen gibt, welche die Vorstellungsfähigkeit auch maßgeblich beeinflussen können.

Dahingegen argumentieren Roberts, Callow, Hardy, Markland und Bringer (2008, S. 201), dass ein solcher Zusammenhang bei motorischen Aufgaben durchaus besteht, weil die Vorstellungsklarheit ein Indikator für die Bewegungsrepräsentation im Arbeitsgedächtnis sein kann. Sie kritisieren aber, dass der VMIQ die Testpersonen auffordert, sich Bewegungen aus der Außenperspektive anhand eines fremden Athleten vorzustellen, wohingegen die visuellen Eindrücke laut Callow und Hardy (2004, S. 174) besser mit der kinästhetischen Vorstellung zusammen passen würden, wenn man sich die Bewegung so vorstellt, als würde man sich selbst von außen beim Bewegungsvollzug beobachten (von Roberts et al. (2008, S. 202) als third-person imagery of the self oder mit dem Akronym EVI (external visual imagery) bezeichnet). Aus diesem Grund nehmen Roberts et al. (2008) eine Modifikation des VMIQ als **VMIQ-2** vor, der im Vergleich zum Originaltest zum einen psychometrischen Validitätskriterien entsprechen soll, zum anderen die Modalitäten kinästhetische Vorstellungsfähigkeit und interne visuelle Vorstellungsfähigkeit als zwei eigenständige Größen behandelt und zuletzt die externe visuelle Vorstellungsfähigkeit anhand des Abbilds der eigenen Person einfordert. Die Vorstellungsaufgaben werden dabei wie folgt formuliert: 1. IVI = „you are looking out through your own eyes“, 2. EVI = „you are watching yourself performing the movement“ und 3. KIN „imagine feeling yourself doing the movement“. Auf einer fünfstufigen Likert-Skala mussten die Probanden die oben genannten 24 Items in

den jeweiligen drei Vorstellungskategorien einschätzen, wodurch sich insgesamt 72 Antworten ergaben. Nach einer CFA-Revision der 24 Items wurden 12 Indikatoren als unzureichend identifiziert und die restlichen 12 Items zu einem 12-item VMIQ-2 mit folgenden Bewegungskategorien zusammengefasst: 1. Walking, 2. Running, 3. Kicking a stone, 4. Bending to pick up a coin, 5. Running up stairs, 6. Jumping sideways, 7. Throwing a stone into water, 8. Kicking a ball in the air, 9. Running downhill, 10. Riding a bike, 11. Swinging on a rope und 12. Jumping off a high wall (Roberts et al., 2008, S. 204f.).

Eine Forschergruppe um Hall entwickelte auf der Basis des MIQ eine ganze Reihe von Fragebögen, um die Qualität von Vorstellungsprozessen in verschiedenen sportlichen Situationen zu messen: den **Imagery Use Questionnaire** (Hall, Rodgers & Barr, 1990), den **Sport Imagery Questionnaire** (Hall, Mack, Paivio & Hausenblas, 1998) und den **Exercise Imagery Questionnaire** (Hausenblas, Hall, Rodgers & Monroe, 1999). Letzterer Fragebogen, der EIQ-AV in der Aerobic-Version, untersucht beispielsweise, ob Sportlerinnen und Sportler in den Kategorien energy (When I imaging exercising, it relieves my stress), appearance (I imagine becoming healthy by exercising) und technique (When learning a new routine, I imagine doing the moves perfectly) Vorstellungsstrategien einsetzen. Die Fragen sind aber eher auf motivationale Aspekte und die grundsätzlichen Modalitäten von Vorstellungsprozessen im sportlichen Übungsprozess gerichtet, so dass die Anwendbarkeit etwaiger Ergebnisse für diese Studie begrenzt ist.

2.3.1.4.2 Mentale Chronometrie

Eine ganz andere Art der Operationalisierung der Bewegungsvorstellung stellt die sog. mentale Chronometrie dar. Munzert (2001b) propagiert aus seinen Überlegungen heraus den Faktor Zeit bzw. den zeitlichen Ablauf als übergreifendes Merkmal der einzelnen Erfassungsebenen. Zeit spielt sowohl im subjektiven Erleben als auch im Verhalten und in neurophysiologischen Prozessen eine Rolle, wenn beispielsweise bestimmt oder eingeschätzt wird, wann bestimmte Vorstellungsmerkmale auftreten oder enden. Im Rahmen der **mentalen Chronometrie** wird die zeitliche Übereinstimmung von mentalem und aktiv ausgeführtem Ablauf verglichen und daraus Rückschlüsse gezogen auf den Zusammenhang zwischen Bewegungsvorstellung, Bewegungsvorbereitung und Bewegungsprogrammierung. Decety et al. (1989) war der erste, welcher derartige Experimente anhand der mentalen und realen Bewältigung verschiedener Gehstrecken (5 m, 10 m und 15 m) durchführte und herausfand, dass es eine zeitliche Übereinstimmung zwischen der Bewegungsvorstellung und der Bewegungsausführung gibt. Die Veränderung der Fortbewegungsbedingungen hat dabei unterschiedlichen Einfluss auf die Übereinstimmung: die Mitnahme eines 25-kg-Gewichts beispielsweise lässt die mentale Bewegungsausführung länger erscheinen als die reale Bewegung (Cerritelli, Maruff, Wilson & Currie, 2000), während aufwärts oder abwärtsführende Strecken wiederum recht exakt mental simuliert werden können (Courtine, Papaxanthis, Gentili & Pozzo, 2004).

Als Beispiele für solche Messverfahren kann der Time-Dependent-Motor-Imagery-screening-Test sowie der Timed-Up-and-Go-Test genannt werden. Jedoch muss erwähnt werden, dass bei dieser Methode vor allem Unklarheit darüber herrscht, ob die Versuchspersonen in der mentalen Bedingung den Zeitbedarf einfach nur schätzen, ohne ihn wirklich gedanklich ablaufen zu lassen (Schott, 2013,

S. 664f.).²² Zu den subjektiv rückgemeldeten Erfahrungen einer Versuchsperson über Anfang, Ende und gegebenenfalls Zwischenstationen des Vorstellungsablaufs oder objektiv messbaren Zeitintervalle derselben können jedenfalls dann neurophysiologische oder peripher-physiologische Parameter in Beziehung gesetzt werden.²³

So kam Parsons (1994, S. 725–727) zu dem Ergebnis, dass Bewegungsvorstellungen stark mit den Bewegungsprogrammen korrespondieren, denn je sicherer und gewohnter eine Bewegung beherrscht wird, desto geringer ist die Abweichung des mental nachvollzogenen Bewegungsablaufs im Vergleich zu seiner realen Ausführung. Eine Bewegungsvorstellung, die dem Motorikprogramm zuwiderläuft oder deren Ausführung motorisch noch wenig automatisiert abläuft, ist zwar möglich, verursacht aber einen zusätzlichen mentalen Aufwand, der auch zeitlich messbar wird. Die Aufgabenschwierigkeit (wie sie in Fitts' law²⁴ konzeptualisiert wird) hat somit die gleichen Auswirkungen auf die Bewegungsvorstellung wie auf die Bewegungsausführung. Bakker, de Lange, Stevens, Toni und Bloem (2007, S. 498) variierten die Schwierigkeit beim Vorstellen von Gehaufgaben beispielsweise nicht nur über die zurückzulegende Strecke, sondern auch über die Breite der Unterstützungsfläche oder das Durchschreiten unterschiedlicher breiter Durchgänge. Bei technisch oder taktisch anspruchsvollen Bewegungen wie das Spielen von Badmintonrallies (Munzert, 2008), das Ausführen von Bodenturnübungen (Calmels & Fournier, 2001) oder das Wasserkunstspringen (Reed, 2002) zeigt sich ebenso eine zeitliche Dissoziation zwischen realer und mentaler Bewegungsausführung. Diese Abweichungen werden damit begründet, dass die enge Beziehung zwischen der Bewegungsausführung und der Bewegungsvorstellung durch verschiedene Faktoren gestört wird. Munzert (2008, S. 33-36) erklärt dies damit, dass bei open-skill-Aufgaben wie einem Ballwechsel im Badminton die Komponente Zeit überhaupt nicht als relevante Größe wahrgenommen wird, weil ein Ballwechsel maßgeblich vom Reagieren auf die gegnerischen Schläge geprägt ist und somit jede Spielsituation neue und andere Anforderungen stellt, weshalb die Dauer vorgestellter Ballwechsel im Vergleich zu den real gespielten verkürzt ist. Calmels und Fournier (2001, S. 147) stellen bei den gymnastischen Turnübungen ebenfalls fest, dass der mental ablaufende Verlauf der Kür schneller von statten geht als der reale, und begründen dies damit, dass der mentale Vorstellungsprozess oft kurz vor dem Wettkampf nochmals zur Einstimmung oder Beruhigung abläuft und daher zeitlich komprimiert wird. Reed (2002, S. 1175) stellt zusätzlich heraus, dass die Vorstellungsdauer im Vergleich zur Realausführungszeit steigt, je komplexer sich die Bewegungsaufgabe darstellt, jedoch sich auch das steigende Expertiselevel des entsprechenden Sportlers positiv auf die mentale Chronometrie auswirkt. Außerdem hat das allgemeine räumliche Vorstellungsvermögen sowie die Fähigkeit zur mentalen Rotation keine Auswirkungen auf das zeitliche Verhältnis von Vorstellungsdauer und Handlungsvollzug. Die Vermutung, dass überlappende mentale Repräsentationen für

²² Sharma, Pomeroy und Baron (2006) haben in ihrer Untersuchung versucht, diese Bedenken durch ein ausgeklügeltes Protokoll zu vermeiden, und bestätigten ebenfalls die Ergebnisse der mentalen Chronometrie.

²³ In der allgemeinen Vorstellungsforschung wurden zur Untersuchung der mentalen Chronometrie vor allem Aufgaben zur mentalen Rotation herangezogen. Der Zeitbedarf für eine mentale Rotation stand dabei oft proportional zum Winkel, um den das mentale Objekt gedreht werden musste. Je komplexer das Rotationsobjekt bzw. je ungewohnter die Bewegungsmöglichkeiten wurden, desto stärker änderte sich auch der Zeitbedarf für die mentale Rotation. Wenn das Rotationsobjekt mehrere Gelenke umfasste oder Bewegungen mit maximaler Gelenkausnutzung verlangt wurden (z. B. bei Bewegungen von Körperumrissen oder Handpositionen), wurden die Abweichungen vom realen Zeitbedarf signifikant höher (Parsons (1987), Parsons (1994)).

²⁴ Fitts (1954) fand in einer experimentellen Studie heraus, dass es einen logarithmischen Zusammenhang gibt zwischen der Schwierigkeit einer Bewegung sowie der Geschwindigkeit, mit der die entsprechende Bewegung ausgeführt werden kann. Je kleiner beispielsweise eine Zieltrefffläche ist, desto schwieriger wird die Bewegungsaufgabe und desto langsamer die Bewegungsausführung. Die Zeit, die eine Person braucht, um eine solche Bewegungsaufgabe gedanklich zu vollziehen, korrespondiert dabei mit der eigentlichen Ausführungszeit und kann somit ein Gradmesser für die Qualität des tatsächlichen Bewegungsvollzugs sein.

den mentalen wie realen Bewegungsvollzug vorliegen, wird dadurch aber nicht in Frage gestellt (Munzert et al., 2009, S. 310). Auch bei Bewegungen mit kurzer Ausführungszeit fanden (Orliaguet und Coello (1998)) am Beispiel des Puttens beim Golf heraus, dass die gedanklich vollzogene Bewegung (im Durchschnitt 467 ms) länger dauerte als der real ausgeführte Putt, welcher unabhängig von der zu überbrückenden Entfernung zum Loch und trotz größerer Bewegungsamplitude im Durchschnitt gleichbleibend nur 262 ms dauerte. Die Autoren schlussfolgern daher bezüglich des Fehlens einer zeitlichen Isochronie: „actual and imagined movement did not share the same information processing. [...] The production of imagined movement would be closer to a perceptive task than to a motor task“ (Orliaguet & Coello, 1998, S. 165f.).

Wenn Bewegungsvorstellungen bewusst in Zeitlupe ablaufen, kann dies unter Umständen negative Folgen für die Ausführungsqualität haben, weil dadurch unterschiedliche neuromuskuläre Muster geschaffen werden. Daher soll mit verlangsamten mentalen Bewegungsabläufen vorwiegend am Anfang des Vorstellungstrainings gearbeitet werden, um vor allem die Bewegungsmerkmale zu memorisieren, ehe dann als letzte Stufe die Originalgeschwindigkeit der Bewegung mental simuliert werden können sollte (Rushal & Lippman, 1998, S. 63). Ebenso kann auch das zeitliche Verhältnis von Bewegungsphasen zueinander untersucht werden. Auch wenn die Bewegung nicht in Echtzeit gedanklich vollzogen werden kann, so kann eine angemessene zeitliche Relation zwischen verschiedenen Bewegungsphasen (z. B. Übungsteilen einer mehrteiligen Bodenturnübung, wie sie (Calmels & Fournier, 2001, S. 145) als sog. Stages beschrieben haben) auf eine zumindest temporal adäquate Vorstellungsfähigkeit hinweisen. Zusammenfassend könnte die Schlussfolgerung stehen, dass eine geringe zeitliche Abweichung der vorgestellten Bewegung vom realen Bewegungsablauf auf eine gut ausgeprägte Bewegungsvorstellung hindeuten kann.

Guillot und Collet (2005, S. 17f.) betonen, dass die zeitliche Vorstellungsdauer einer Bewegung auch länger wird, wenn Athleten sich in der konkreten Wettkampfvorbereitung befinden oder nicht den ganzen Bewegungsablauf gedanklich ablaufen lassen, sondern auf bestimmte Bewegungsaspekte ihre Aufmerksamkeit lenken. Ebenso verlängert sich die Vorstellungsdauer, je komplexer das koordinative Anforderungsprofil einer Bewegung ist, wie Guillot und Collet (2005, S. 14f.) dies für die Sportarten Tennis (Aufschlag und Volley) sowie Bodenturnen (Radwende-Flickflack-Salto rückwärts) nachweisen konnte. Munroe, Giacobbi, Hall und Weinberg (2000, S. 133f.) erklären, dass Probanden in Trainings- oder Testsituationen ohne Zeitdruck mehr Zeit für die Ausbildung einer präzisen Bewegungsrepräsentation verwenden. Außerdem scheint bei komplexen Bewegungsaufgaben die Vorstellungsgenauigkeit wichtiger zu sein als zeitliche Bewegungsmerkmale. Guillot und Collet (2005, S. 14f.) halten in diesem Zusammenhang fest, dass beim Vorstellen der eigenen Bewegung aus der Fremdperspektive die Bewegungsdauer keine dominante Rolle spielt, weil die Sportler zuerst versuchen, Vorstellungsgenauigkeit zu erreichen.

Zusammenfassend kann mit Martin, Moritz und Hall (1999) festgehalten werden, dass Sportler oder Coaches lernen müssen, Vorstellungsbilder situativ passend einzusetzen (z. B. durch Geschwindigkeitsreduzierung, um ein Bewegungsmerkmal speichern oder noch detaillierter visualisieren zu können). Anweisungen bezüglich der Vorstellungsdauer von Bewegungen müssen so genau formuliert sein, dass die mentalen Verarbeitungsstrategien an die Ziele des Sportlers oder Coaches angepasst werden können.

2.3.1.4.3 Mentale Rotation

Eine andere Möglichkeit besteht darin, die Bewegungsvorstellung computergestützt über die kognitive Leistung, ein Objekt mental zu repräsentieren und durch Rotationen zu transformieren, zu operationalisieren. Beim **Mental Rotation Test (MRT)** nach Vandenberg und Kuse (1978) müssen die Versuchspersonen in zwei Durchgängen jeweils in drei Minuten zehn Rotationsaufgaben bewältigen und sich in Form eines paper-and-pencil-Tests für die richtigen Antwortmöglichkeiten entscheiden. In einer Weiterentwicklung dieser Methode wurde der **Controllability-of-Motor-Imagery-Test** von Nishida, Katube und Inomata (1986) (modifiziert nach Naito (1994)) entwickelt, in dem die Versuchsperson eine Reihe von Instruktionen (pro Instruktion muss ein Körperteil mental in eine neue Position bewegt werden) befolgen und die finale Körperposition des Vorstellungsobjekts reproduzieren muss, ohne sich dabei aktiv zu bewegen. Der Test wurde von Schott (2013) deutschsprachig modifiziert als **Test zur Kontrollierbarkeit der Bewegungsvorstellungsfähigkeit (TKBV)** und umfasst insgesamt 20 Vorstellungsaufgaben. Der Proband muss pro Aufgabe sechs aufeinanderfolgende Instruktionen, die im Abstand von drei Sekunden erfolgen, mit geschlossenen Augen mental manipulieren, wobei pro Instruktion ein Körperteil angesprochen wird. Nach der letzten Anweisung muss die vermutete Endposition entweder anhand von fünf vorgeschlagenen Körperpositionen wiedererkannt (siehe Abb. 26) oder ohne auszuwählende Vorgaben frei wiedergegeben werden, indem die finale Position so schnell wie möglich aktiv eingenommen wird.

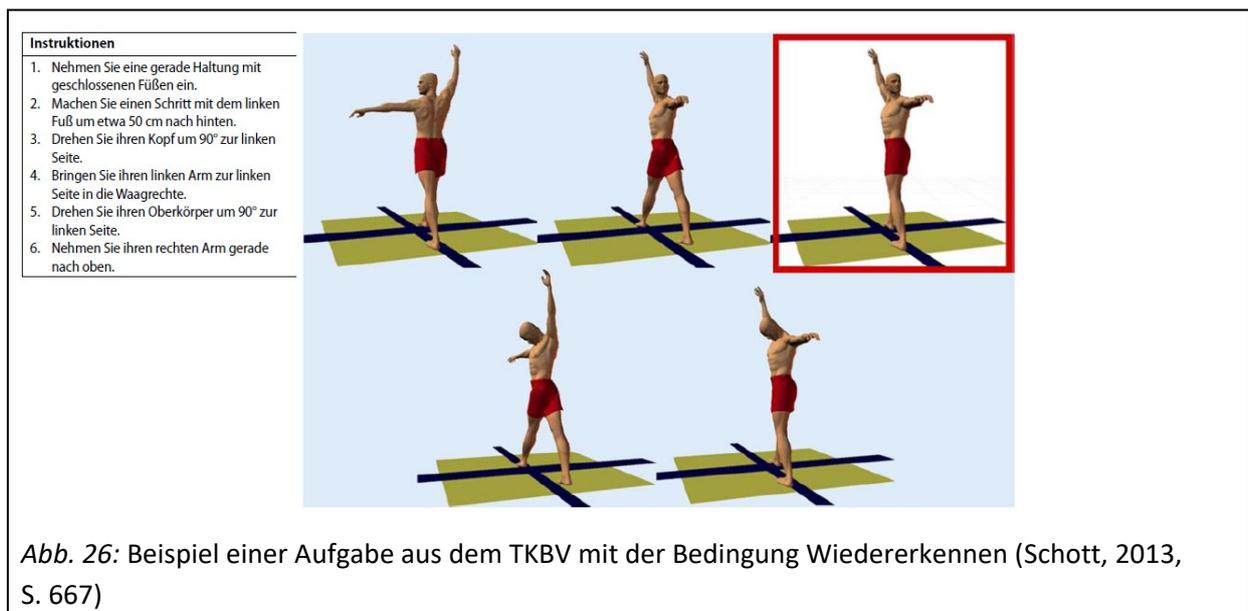


Abb. 26: Beispiel einer Aufgabe aus dem TKBV mit der Bedingung Wiedererkennen (Schott, 2013, S. 667)

2.3.1.4.4 Movement Imagery Specific Test

Ein weiteres Testverfahren stellt der **Movement Imagery Specific Test (MIST)** von Moreau et al. (2010) dar. Das Besondere an diesem Verfahren ist die Tatsache, dass die Aufgaben für konkrete Sportarten (hier Fechten, Judo und Ringen) entwickelt wurden und somit sportartspezifische Vorstellungsfähigkeit verlangt wird. Jeder Test besteht dabei aus 20 verbalen Äußerungen (auf visuelle Informationen wurde bewusst verzichtet), welche eine sportartspezifische Situation beschreiben, deren räumliche Struktur gedanklich transformiert werden muss (siehe Abb. 27). Pro Äußerung muss aus zwei Antwortmöglichkeiten die korrekte ausgewählt werden. Es wurden besonders aussagekräftige und einfache Situationen gewählt, damit Profisportler wie Novizen gleiche Voraussetzungen für die Aufgabenbewältigung

hatten. Bezüglich der Perspektive (innen oder außen) oder der Modalität (visuell oder kinästhetisch) wurden keine besonderen Instruktionen gegeben. Insgesamt hatten die Probanden drei Minuten Zeit, die 20-Item-Aufgabe zu bearbeiten. Anhand des Gesamtscores (maximal 20 Punkte) wurde die Vorstellungsfähigkeit in der jeweiligen Sportart bestimmt.

MIST Fencing

Starting from an engagement in sixte with my opponent whose left hand is armed. If I am performing a lunge, am I in his/her parry of quarte or sixte?

- Quarte
- Sixte

MIST Judo

I am performing *juji-gatame* on my opponent's right arm. Which of my legs is holding his head down?

- Right
- Left

MIST Wrestling

I am performing an *ankle lace* to my opponent's left. Which of his ankles is on the top of the other?

- Right
 - Left
-

Abb. 27: Beispiele für Fragen und Antwortmöglichkeiten des MIST (Moreau et al., 2010, S. 99)

2.3.1.4.5 Zusammenhang verschiedener Testverfahren

Collet, Guillot, Lebon, MacIntyre und Moran (2011) haben in ihrem Ansatz versucht, die bisher isoliert nebeneinanderstehenden psychometrischen Tests, qualitativen Verfahren, chronometrischen Methoden sowie psychophysiologischen Techniken aufeinander zu beziehen und daraus einen gemeinsamen Motor Imagery Index (MII) zu entwickeln. Der Index setzt sich aus verschiedenen Komponenten zusammen, muss aber erst noch empirisch validiert werden:

Self-estimation of imagery quality, psychometric assessment of imagery vividness, MC (estimation of the differences between actual and imagined duration of movement execution), and from three psychophysiological concomitants of imagery (based on electrodermal and cardiac activities). (Collet et al., 2011, S. 91)

Moreau et al. (2010, S. 103–105) stellen nach einem Vergleich der drei Methoden Movement Imagery Questionnaire (MIQ-R), Mental Rotation Test (MRT) und Movement Imagery Specific Test (MIST) fest, dass Selbsteinschätzungsfragebögen (wie der MIQ-R) keine Korrelation aufweisen zu Verfahren, welche die konkrete Bewegungsvorstellungsleistung zu messen versuchen (wie beim MRT oder MIST). Jedoch konnte ein Zusammenhang zwischen dem MIST und dem MRT nachgewiesen werden, was darauf hindeutet, dass beide nach ähnlichen Prozessen ablaufen oder von gemeinsamen Merkmalen beeinflusst werden. Außerdem konnte bei den Verfahren MIST und MRT eindeutig nachgewiesen werden, dass Profisportler gegenüber Novizen eine bessere Vorstellungsfähigkeit erzielen, wohingegen diese Unterscheidung in den MIQ-R-Ergebnissen nicht sichtbar wurde. Insgesamt kann festgehalten werden,

dass die Vorstellungsfähigkeit nicht allein mittels Selbsteinschätzungsaufgaben beurteilt werden kann, sondern mit experimentellen Messungen der räumlichen Vorstellungsfähigkeit mittels „performance-based tests“ (Moreau et al., 2010, S. 103) wie dem MRT oder MIST ergänzt werden sollte. Es soll aber auch nicht verschwiegen werden, dass die beiden genannten Tests die Vorstellungsfähigkeit nur indirekt über korrekte oder inkorrekte Antworten messen. Die kognitiven Prozesse sowie Strategien zur Aufgabenbewältigung bleiben dabei unberücksichtigt. McAvinue und Robertson (2009, S. 115) behaupten, dass die kinästhetische Subskala des MIQ-R eine Kohärenz mit der Messung der mentalen Chronometrie aufweist. Dies lässt vermuten, dass sich diese beiden Kategorien auf die gleichen Fähigkeiten beziehen, während die visuelle Subskala des MIQ sowie die kinästhetische Subskala des QMI scheinbar auf andere Fertigkeiten ansprechen.

Wie oben bereits erwähnt, gehen Dausgs et al. (1989, S. 48) davon aus, dass für die Operationalisierung einer Bewegungsvorstellung die essentiell erscheinenden Bedeutungseinheiten in ihrer Struktur identifiziert und als elementare (sensomotorische) Lehr-Lern-Operationen sprachlich und bildhaft ausgewiesen werden müssen, wobei vor allem didaktisch leicht greifbare Aspekte (z. B. Umkehrstellen und Lageveränderungen von Körperteilen, physikalische Wirkungseffekte, eindeutige Geschwindigkeits- und Richtungsangaben) auszuwählen sind. Die Ausgliederung solcher Teilaspekte der Gesamtbewegung und der damit zu identifizierenden kognitiven Bedeutungseinheiten (sowie die Frage, welche sensomotorischen Operationen als wesentlich und elementar zu betrachten sind) hängt dabei von der angestrebten didaktischen Anforderungssituation ab. Zwar sind die sensomotorischen Operationen als einer der fertigen Bewegungsvorstellung zugehörigen regulativen Teileinheit keinesfalls identisch mit Lehr-Lern-Operationen, welche als Informationsschritt auf dem Weg zur Ansteuerung einer Teilbewegung aufgefasst werden können, stellen aber in ihren sicht- oder messbaren Korrelaten die Grundlage zur Realisierungsbewertung dar. Voraussetzung dafür ist, dass ein Lehrender diese Operationen unter Beurteilungs- und Korrekturgesichtspunkten mit Erfolg als intendierte Verhaltenseinheiten des Lernenden wahrnehmen und identifizieren kann (Blischke, 1986, S. 13).

Zur Registrierung der Bewegungsvorstellung am Beispiel der großen Körperwelle vorwärts wurde von Dausgs et al. (1989, S. 48–51) ein Bildkarten-Auswahl-Test (BKAT) entwickelt, in welchem verschiedene Einzelbilder (siehe Abb. 28) entsprechend des korrekten Bewegungsablaufs in die richtige Reihenfolge (in Form eines Reihenbildes) gebracht werden müssen. Es wird dabei davon ausgegangen, dass richtige Auswahlentscheidungen auf „eine bewußte [sic!] Vorstellung von den für die Zielbewegung essentiellen sensomotorischen Operationen und ihrer räumlich-zeitlichen Struktur“ (Dausgs et al., 1989, S. 49) hindeuten.

Der Test vereinigt in sich einerseits eine Memorations- bzw. Rekognitionsleistung (Identifizieren der in der Lehrinformation kodierten ‚zentralen Knotenpunkte‘ der Bewegung), andererseits eine kognitive Konstruktionsleistung (Erschließen von in der Lehrinformation nicht explizit enthaltenen ‚Zwischenkonfigurationen‘ der Bewegung anhand der bereits entwickelten Bewegungsvorstellung und Arrangieren aller Konfigurationen entsprechend ihrer zeitlichen Abfolge. (Dausgs et al., 1989, S. 49)

Wenn nun also ein Proband aus den verschiedenen Bildstapeln die richtige Körperposition im Zeitverlauf der Bewegung wählt bzw. falsche Körperpositionen ausschließt, steigt seine individuelle Testleistung oder verringert sich analog, wenn eine fehlerhafte Auswahl getroffen wird.

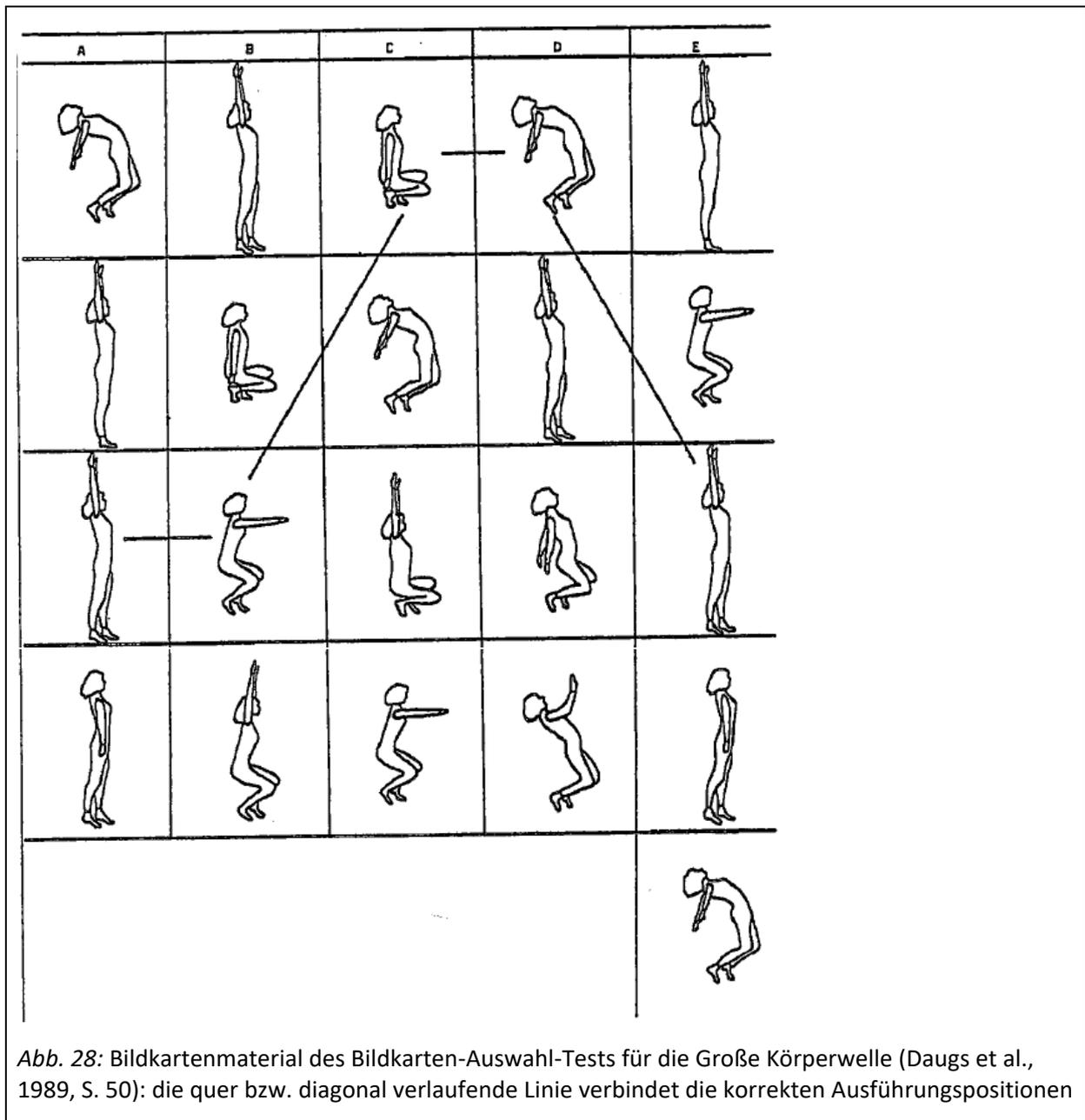


Abb. 28: Bildkartenmaterial des Bildkarten-Auswahl-Tests für die Große Körperwelle (Daugs et al., 1989, S. 50): die quer bzw. diagonal verlaufende Linie verbindet die korrekten Ausführungspositionen

Blichke (1986, S. 11f.) stellt fest, dass eine Operationalisierung von Bewegungsvorstellungen auf den sensorischen Output, also die Bewegungsausführung verwiesen ist, weil die Struktur des konzeptuellen Gedächtnisses uns nicht unmittelbar zugänglich ist. Ein gezeigter Bewegungsablauf stellt somit die Konkretion der jeweils vorhandenen Bewegungsvorstellung dar und lässt auf deren Qualität und Ausprägung schließen. Um die Bewegungsvorstellung identifizieren zu können, müssen die essentiell erscheinenden Bedeutungseinheiten der Bewegung in ihrer Struktur operational repräsentiert werden (**Sollwertfestlegung**), wobei sich besonders solche Aspekte der Kriteriumsbeziehung anbieten, welche lernrelevante und didaktische Bedeutung haben.

Dies können z. B. Teilbewegungen sein [...], die sich im Hinblick auf die endgültige Zielansteuerung (Resultatoptimierung, z. B. Kugelstoß, bzw. Sollwertentsprechung, z. B. Ballett-Position) als essentiell oder als korrekturrelevant erwiesen haben, und deren wesentliche dynamischen und funktionellen Charakteristika und wechselseitigen Relationen ggf. durch Geschwindigkeits- und Richtungsangaben, durch Kennzeichnung von Ortsverhältnissen, von Umkehrstellen und Lageveränderungen, aber auch

von räumlich-zeitlichen Relationen und physikalischen und/oder physiologischen Wirkungseffekten um Ausdruck gebracht werden können, wo immer dies notwendig erscheint. (Blichke, 1986, S. 12)

Daher soll im nun folgenden Kapitel näher auf die Festlegung und Präzision von Sollwertformulierungen sowie die Fähigkeit zum Analogie-Schluss eingegangen werden, um die allgemeinen und sehr theoretischen Ausführungen über das Phänomen der Bewegungsvorstellung auf die in dieser Studie zu untersuchende Wurfbewegung zu beziehen.

2.3.2 Sollwertpräzision und Analogie-Schluss

Als Ergebnis einer theoretischen Beschäftigung mit einer Bewegung steht am Ende die Formulierung eines Sollwertes, der die optimale Bewegungsausführung beschreibt. Die für diese Studie notwendige Sollwertformulierung wird für die leichtathletische Schlagwurfbewegung in den Kapiteln 4.2.1.1 (Wissen über den räumlich-zeitlichen Verlauf) und 4.2.2.1 (biomechanisches Wissen) ausgeführt. Mit dem Begriff **Sollwertpräzision** ist dann die Fähigkeit gemeint, sich eine Bewegung genau vorzustellen und rekonstruieren sowie antizipieren zu können, auf welche Informationen man bei einer effektiven Wahrnehmung stoßen könnte. Neumaier (1988, S. 283) führt für dieses Phänomen synonym den Begriff **Vorinformation** ein und meint damit generell „alle dem Beurteiler verfügbaren Informationen über die zu beurteilende Bewegung [...], die zur Entwicklung antizipierender Schemata im Wahrnehmungsprozeß [sic!] beitragen können.“ Daus und Blichke (1984, S. 388) sprechen in diesem Zusammenhang von sensomotorischer Konzeptualisierung, wenn mittels sprachlich gestützter Bewegungsbilder eine Bewegungsvorstellung nach räumlichen, zeitlichen und dynamischen Gesichtspunkten initiiert und die Aufmerksamkeit auf die Auswahl von Knotenpunkten der Bewegung, ausgezeichneten Lagen oder sensomotorischen Operationen gelegt wird. Diese Sollwertpräzision der Testpersonen soll mit einem Fachwissenstest (siehe Kapitel 4.2) bestimmt werden, um deren Einfluss auf die Beurteilungsleistung zu ermitteln. Mit Morrison (2000, S. 22) ist allerdings auch festzuhalten, dass die Art, wie ein Beobachter eine erfolgreiche Bewegungsausführung definiert und analysiert, und das verwendete Format der Bewegungsbeschreibung unterschiedlich sein können und die Analysefähigkeit beeinflussen.

Da nicht alle Körperteilbewegungen synchron-optisch erfasst werden können, muss aus dem Bewegungsverhalten foveal fixierter Körperteile auf das nicht-beobachtete Verhalten anderer Körperteile geschlossen werden können (Analogie-Fähigkeit). In einem Ist-Sollwert-Vergleich kann dann eine beobachtete Bewegung mit dieser optimalen Bewegungsvorstellung verglichen werden, um Übereinstimmungen oder Abweichungen davon feststellen zu können.²⁵ Radford (1990, S. 11) präzisiert: „An observation is an outcome of the comparison of the stimuli gathered from a display to a generic representation of movement stored within the observer’s memory template.“

Der Begriff der **Bewegungsvorstellung** (siehe Kapitel 2.3.1) liefert hilfreiche Informationen für ein besseres Verständnis der Sollwertpräzision. Grundsätzlich geht es um eine möglichst realistische Bewe-

²⁵ Barrett (1984, S. 24) weist darauf hin, dass für eine fundierte Bewegungsbeobachtung neben der Sollwert-Analyse auch die Umweltbedingungen, in der die Bewegung stattfindet, präzise antizipiert und wahrgenommen werden müssen.

gungsvorstellung (template formation), die als interne Repräsentation eines oder mehrerer, im Langzeitgedächtnis gespeicherter menschlicher Bewegungsmuster (nicht misszuverstehen als Anzahl isolierter Einzelbilder) verstanden werden, welche zum Zwecke eines Ist-Sollwert-Vergleichs abgerufen werden. Die Bewegungsvorstellung setzt sich dabei aus deklarativem und prozeduralem Wissen sowie affektiven Eindrücken zusammen. Während sich deklaratives Wissen hauptsächlich aus theoretischen Fakten zusammensetzt (zu beschreiben als *wissen, dass*) und als semantische Information zu verstehen ist, zielt das prozedurale Wissen auf die Bewegungshandlung selbst ab (verstanden als *wissen, wie*) und greift überwiegend auf optische und akustische Eindrücke zurück. Daher ist es für Sportler oft auch schwierig, verbal zu beschreiben, wie sie eine Bewegung ausführen, weil diese hauptsächlich prozedural abgebildet ist. Deklarative Informationen müssen demnach mehrschichtig mit prozeduralem Wissen verknüpft werden, damit sie für die Bewegungsbeobachtung nutzbar gemacht werden können.²⁶ Durch Beobachtungserfahrung wird die Bewegungsvorstellung erweitert und so optimiert, dass für das Erkennen bewegungsrelevanter Informationen weniger Zeit benötigt wird, was vor allem für schnell ablaufende Bewegungen vorteilhaft ist (Radford, 1990, S. 8–11).

Generalisierung und Gruppierung stellen zwei wichtige Faktoren dar, um die Bewegungsvorstellung für die Bewegungsbeobachtung nutzbar zu machen. Durch Generalisierungen ist die Bewegungsvorstellung auf verschiedene Situationen und ähnliche Bewegungen übertragbar, wie dies beispielsweise bei der Übertragung von Technikkriterien des leichtathletischen Schlagwurfs auf den Handballschlagwurf oder den Volleyballschmetterschlag der Fall ist. Mittels Gruppierung von verschiedenen Technikdetails zu einer Beobachtungseinheit wird die Beobachtung vereinfacht und der Fokus kann auf bestimmte Bewegungsmerkmale gelegt werden, ohne die Gesamtbewegung aus den Augen zu verlieren (Radford, 1990, S. 18–20). Aus dieser Sicht scheint auch die in dieser Studie bevorzugte Verwendung der Meinel'schen Bewegungsmerkmale als bewährte Gruppierungshilfe gerechtfertigt, um den Beurteilungsprozess so zu vereinfachen, dass auch die nahezu gleichzeitige Wahrnehmung mehrerer Bewegungsmerkmale gelingen kann. Die prozedurale Bewegungsvorstellung lässt sich dabei aufgliedern in einen eigenrealisationsfreien Teil, der auf visuellen und verbalen Elementen beruht, und einem eigenrealisationsgebundenen Teil²⁷, der mit kinästhetischen, akustischen, vestibulären und taktilen Reafferenzen gekoppelt ist (Tidow, 1981, S. 264). Die Beschäftigung mit der Sollwertpräzision erfolgt in diesem Kapitel aus einer eigenrealisationsfreien Perspektive, während dann im nächsten Kapitel (0) der Frage nachgegangen wird, welche Bedeutung die Bewegungserfahrung aus der Eigenrealisation für die Bewegungsbeurteilungsleistung eines Sportlehrers oder Trainers hat.

2.3.2.1 Empfehlungen für die Formulierung von critical features

Knudson (2013, S. 76) weist in seinem QMD-Modell (Qualitative Movement Diagnosis) darauf hin, dass in der Vorbereitung auf eine Bewegungsbeobachtung die intensive Beschäftigung mit der Bewegung selbst unumgänglich ist. Um Bewegungen überhaupt angemessen beurteilen zu können, ist ein möglichst umfassendes Vorwissen (prerequisite knowledge) über die optimale Ausführungsqualität (acti-

²⁶ Helsen und Pauwels (1993, S. 384) kommen in einer Untersuchung über Sehstrategien von Fußballern in taktischen Entscheidungssituationen zu folgender Einschätzung: „A greater amount of specific declarative knowledge [...] does not guarantee better performance.“

²⁷ Radford (1990, p.9) unterscheidet in ähnlicher Form „knowledge formulated from watching others perform“ und „knowledge gained from learning how to perform“

vity or movement) notwendig, welches unter anderem durch die Beschäftigung mit fachwissenschaftlicher Literatur erworben werden kann. Dabei ist die Bewegungsvorstellung geprägt von Schlüsselmerkmalen, sog. „critical features“ (Barrett, 1984, S. 25), die bei einer gelungenen Bewegungsausführung beobachtbar sind, weil sie unabhängig von der individuellen Technik eines Sportlers immer zu einem bestimmten Zeitpunkt (temporal) auf sehr ähnliche Weise (spatial, d. h. räumlich) ausgeführt werden und den Verlauf der Bewegung maßgeblich beeinflussen.²⁸ Bevor man sich daher an konkrete Sollwertformulierungen für die leichtathletische Schlagwurfbewegung macht, sollen die Rahmenbedingungen für die Festlegung passender critical features geklärt werden.

Knudson (2013, S. 81f.) empfiehlt, eine Hierarchie dieser **critical features** festzulegen, d. h. welches Merkmal ist wichtiger für das Gelingen der Gesamtbewegung als ein anderes, und die Aufmerksamkeit bei der Beobachtung dann auch auf diese priorisierten Merkmale zu richten. Auch der Zeitpunkt innerhalb der Bewegungssequenz, zu dem das entsprechende Merkmal ausgeführt werden soll, muss möglichst exakt bestimmt werden.²⁹ Je umfassender das Bewegungswissen eines Beobachters ausgeprägt ist, desto besser wird ihm die Identifizierung und Hierarchisierung der Schlüsselemente gelingen, auch wenn sich diese beispielsweise nicht immer in klar definierten biomechanischen Parametern ausdrücken, sondern oft nur anhand ihrer Auswirkungen auf die Bewegung beobachtet werden können.

Dass der Fokus bei der Beobachtung in der Lehrpraxis häufig auf die falschen Schwerpunkte gelegt wird, kann nach Barrett (1984, S. 25) zwei Gründe haben: (a) fehlende Kenntnis über den korrekten Bewegungsablauf oder (b) unterschiedliche Vorstellungen über Bewegungsrealisierung und Fertigungsaneignung im Allgemeinen. Dabei dürfte auch eine Rolle spielen, über wieviel Beobachtungserfahrung ein Beurteilender verfügt, weil anzunehmen ist, dass der wiederholte, langjährige Umgang mit einem Beobachtungsgegenstand zu einer Verbesserung des visuellen Wahrnehmungsprozesses führt. Es ist jedoch zu bedenken, dass dieses erfahrungsbasierte Wissen meist spezifisch auf die entsprechende Sportart ausgerichtet ist und eine hohe Praxistauglichkeit besitzt, jedoch nicht generalisierbar ist auf beliebige andere Sportarten (Knudson, 2013, S. 77).

Die critical features sollten dabei so präzise wie möglich definiert werden und gleichzeitig ein gewisser Toleranzbereich (**range of correctness**) eingeräumt werden. Dieser Toleranzbereich versucht der Tatsache Rechnung zu tragen, dass es große interindividuelle Unterschiede in der technischen Ausführung gibt, ohne dass sich dies gleich dramatisch auf die sportliche Leistung auswirken muss. So kann beispielsweise die Beinstellung bei der Ausführung eines scharfen Baseballwurfes je nach individueller Technikausprägung sehr unterschiedlich aussehen und es kann hilfreich sein für die Bewegungsbeobachtung, wenn ein Toleranzbereich festgelegt wird, innerhalb dessen eine effektive Bewegungsausführung stattfinden kann (z. B. zwischen 50 und 90 Prozent der Körperlänge des Werfers). Killing (2009, S. 154f.) erklärt, dass der leichtathletische Hochsprung eine Disziplin mit vielen Technikspielräumen ist und deshalb viele interindividuelle Varianten in einzelnen Bewegungsphasen möglich sind, ohne große Leistungseinbußen befürchten zu müssen. Die Bestimmung eines solchen Toleranzbereichs wird jedoch immer dann problematisch, wenn die Expertenmeinung von erfahrenen Trainern, die auf Grundlage von best-practice-Beispielen beurteilen, nicht mit biomechanischen Theorien über die optimale

²⁸ Critical features können in Form von komplexen Ideen, subtilen Punkten, allgemeinen Wissensinformationen, professionellen Sprachinhalten (Jargons) oder präzise bestimmbar Werten auftreten (Knudson (2013, S. 85)).

²⁹ Knudson (2013, S. 86) kritisiert, dass es an konkreten Taxonomien für die critical features in verschiedensten Sportarten fehle: „Some books provide general technique points and cues for teaching motor skills, but often these are just recent versions of old opinion-based instructional books. It would be most desirable to get experts together with scholars from many subdisciplines of kinesiology and sport coaches to create universal or official taxonomies of critical features for the skills and fundamental movement patterns of specific sports.“

Bewegungsausführung übereinstimmt. Gerade auch im Schulsport kann die Lösung darin bestehen, Techniksollwerte nicht zu eng festzulegen, sondern einen gewissen Spielraum für mögliche generalisierte Ausführungsvarianten zu lassen (Radford, 1990, S. 11f.).

QMD is easier and more reliable if the professional can establish a range of correctness for the critical features and technique points of a movement. This is often not easy because desirable and optimal movement patterns have not been established. (Knudson, 2013, S. 85)

Grundsätzlich können die Schlüsselemente auf die Bewegungsausführung selbst bezogen werden (knowledge of performance), aber auch auf das Bewegungsziel ausgerichtet sein (knowledge of results). Beispielsweise kann die Qualität des leichtathletischen Schlagwurfs bewegungsbezogen an der richtigen Kopplung der Körperteilbewegungen festgemacht oder auch resultatsbezogen an der Ballflugbahn (Wurfweite und Abwurfwinkel des Wurfgeräts) beurteilt werden (Knudson, 2013, S. 84f.). In dieser Studie geht es aber ausschließlich um eine bewegungsbezogene Beurteilung, weil die Ballflugbahn filmisch nicht zufriedenstellend erfasst werden kann.

2.3.2.2 Technik-Cues

Da die critical features sehr technisch formuliert sind, finden sich in manchen Bewegungsbeschreibungen sog. **Technik-Cues**, welche die Technikkriterien in eine dem Sportler verständlichere Sprache zu übersetzen versuchen:

A cue is defined as a guiding suggestion or a stimulus that excites the imagination to action. Cues are short, catchy phrases that call the student's attention to key components of a skill. A cue projects a clear description of a skill component into the student's mind. (Fronske, 2014, S. 2)

Cues als kurze verbale Erinnerungen können dem Sportlehrer oder Trainer helfen, in der komplexen Beobachtungssituation (z. B. im Sportunterricht) den Überblick über die critical features zu bewahren und diese in konkrete Handlungsanweisungen zu übersetzen. Mittels Cues können komplizierte Bewegungsinformationen komprimiert werden, die dann bei der Bewegungsbeobachtung unbewusst wieder ins Gedächtnis gerufen werden. Um bei der Beurteilungsaufgabe nicht überfordert zu werden (beispielsweise durch zu schwierige synchron-optische Wahrnehmung), soll in einem Bewegungsanalyse-durchgang nur eine begrenzte und vorher festgelegte Anzahl an Cues beobachtet werden. So kann die Lehrkraft gezielt ihren Fokus auf eine besondere Komponente der Bewegungsausführung lenken (Fronske, 2014, S. 3). Es sei nochmals betont, dass es für eine Bewegungsbeurteilung nicht ausreicht, nur die Cues zu kennen, sondern auch die darin beinhalteten Detailinformationen über die richtige Bewegungsausführung verstanden zu haben.

Der besseren Merkfähigkeit wegen soll ein Cue nicht mehr als sechs Wörter umfassen, wobei dieser gegebenenfalls noch mit mehr Detailinformationen angereichert werden kann. Nach Dunham (1994) eignen sich dabei bildhafte Umschreibungen der richtigen Technik (z. B. ‚Jucke Dich mit dem Tennisschläger am Rücken‘ als Bewegungsanweisung für die Ausholbewegung des Tennisschlägers beim Aufschlag) besser als technisch exakt formulierte Instruktionen, auch wenn diese den Bewegungsablauf eigentlich genauer beschreiben. Auch die Benutzung von Cues oder einfachen Wendungen als Vereinfachung komplexerer Bewegungsabläufe (z. B. ‚rechts – links‘ als Beschreibung der richtigen Schritt-

folge beim Basketballkorbleger einer Rechtswerfers) kann hilfreich sein für die Bewegungsbeobachtung, wobei solche Verbalisierungen in einem verbalen Pre-Training eingeübt und dem Beobachter verständlich gemacht werden müssen. Diese Vereinfachung hilft beim schnellen Analysieren von Hauptfehlern und deren Ursachen, weil die Bewegungsausführung unmittelbar mit den Cues verglichen werden kann (z. B. ist der Cue ‚Schwing den Schläger von unten nach oben‘ einfach zu beobachten) (Fronske, 2014, S. 12–14). Morrison und Reeve (1993) schlagen vor, Bewegungs-Cues nach den Kategorien action, content, qualifications und special conditions viergliedrig zu gestalten. Zuerst soll ein Verb festgelegt werden, welches die gewünschte Bewegung beschreibt (action), dann wird auf das Körperteil verwiesen, welches die Bewegungshandlung ausführt (content) und abschließend wird spezifiziert, wie der Bewegungserfolg erreicht wird (qualifications). Ganz zuletzt können noch Informationen angefügt werden, die für die Bewertung der Bewegungsausführung noch hilfreich sind (special conditions).³⁰ Solche Cues können dabei für alle Phasen des Bewegungsverlaufs (z. B. Vorbereitungs-, Haupt- oder Endphase) formuliert werden.

There is an important connection between teaching motor skills and later assessing them qualitatively to improve performance. A common language base and understanding of the critical features of a movement make QMD easier and more effective. (Knudson, 2013, S. 88)

Die Cues können auch dazu eingesetzt werden, **Rückmeldung oder Instruktion** über den Bewegungsverlauf in eine Sprache zu fassen, die beim Sportler ankommt, und als gemeinsames Vokabular für Feedbackprozesse dienen. Denn mit Knudson (2013, S. 94) kann übereinstimmend festgestellt werden: „What good is advanced knowledge about human movement if you cannot communicate effectively to your clients?“ Um den Sportler dabei nicht zu überfordern und auf Bekanntes zu verweisen, kann das Feedback über die technische Ausführung derselben Struktur folgen, nach der auch die Bewegungsbeschreibung präsentiert wurde. Die **Technik-Cues** kann der Sportler dann als konkrete Handlungsanweisung praktisch umsetzen.³¹ Je größer dabei das Repertoire einer Lehrperson an solchen cue words und cue phrases ist, desto besser wird es ihm gelingen, dem Sportler entsprechend seines Alters, seiner Bewegungserfahrung und seines Interessensstandes Informationen und Rückmeldungen über die korrekte Bewegungsausführung zu kommunizieren (Knudson, 2013, S. 91).

You must know the correct movement patterns to analyze the skills of your students and give technical feedback. [...] Teaching sport skills is easier and more effective when you focus on one to three cues and then give feedback on those cues. (Fronske, 2014, S. 5)

2.3.2.3 Einsatz von Analysebögen

Analysebögen dienen dazu, die vielen Technikkriterien einer Bewegung in einer Form zusammenzufassen und zu vereinfachen, die eine Bewegungsbeobachtung und Bewegungsbeurteilung möglich machen. Im Rahmen dieser Studie übernehmen sie an zwei Stellen eine wichtige Funktion: 1. als Beurteilungsraster zur Messung für die Beurteilungsleistung (siehe Tab. 53 im Appendix) und 2. als Detailana-

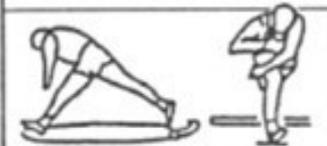
³⁰ „This structure is illustrated in the cue phrase for overarm throwing ‚Swing your arm forward, level to the ground.‘ ‚Swing‘ indicates the action; ‚arm‘ indicates the content; ‚forward‘ qualifies the action; and ‚level to ground‘ is a special condition defining the movement.“ (Knudson (2013, S. 90))

³¹ „Good cue phrases communicate the essence of a critical feature or technique point concisely so that the performer can remember it during practice.“ (Knudson (2013, S. 88))

lysebogen zur Bestimmung der Qualität der Eigenrealisation (siehe Kapitel 4.4 und Tab. 50 im Appendix). Während das Beurteilungsraster so einfach konstruiert ist, dass es bei der morphologischen Bewegungsbeurteilung angewendet werden kann, ist der Detailanalysebogen eher auf die exakte Bestimmung der einzelnen Technikmerkmale ausgerichtet. Trotz ihrer sehr unterschiedlichen Detailliertheit sollen bei der Konstruktion der beiden Analysebögen die nun folgenden Erkenntnisse berücksichtigt werden.

Grundsätzlich wird ein Analysebogen (task sheet) aus den critical features oder den konkreten Bewegungs-Cues entwickelt, in welchem (auch unter Einbeziehung von Photos oder Zeichnungen) exakt beschrieben wird, wie eine Bewegung korrekt ausgeführt wird. Genauso können die gleichen Bögen aber auch für Bewegungsbeurteilung und Feedbackprozesse verwendet werden, damit eine möglichst hohe Übereinstimmung erreicht wird zwischen dem, was gelehrt und dem, was bewertet und rückgemeldet wird. Dies wirkt sich dann positiv auf den Lernerfolg des zu betreuenden Sportlers aus (Knudson, 2013, S. 89–91). Tidow (1983) hat beispielsweise für die azyklische Bewegung der O-Brien-Technik im Kugelstoßen eine sehr detaillierte Bewegungsanalyse erstellt und seine Beobachtungskriterien in ein Analysebogen-Grundmodell gebracht, welches in acht aufeinanderfolgende Phasen strukturiert ist (siehe Abb. 29). Zur Darstellung und Beobachtung jeder dieser Phasen hat er vier Kategorien unterschieden:

1. Visuelle Darstellung wichtigster Bewegungspositionen durch Umrisszeichnungen (Kontogramme)
2. Phase: Sprachliche Kennzeichnung der Phase (z. B. Ausgangsstellung, Standwaage, Starthocke)
3. Bezug: Beobachtungspunkte einer Einzelphase (Phasenelemente) (z. B. Fußstellung, Stoßarm)
4. Kriterium: Beurteilungsmerkmal eines gelungenen Phasenelements (z. B. Fuß in Stoßrichtung, Kugel am Hals)

	PHASE	BEZUG	KRITERIUM
	1. Ausgangs- stellung	1.1 Fußstellung 1.2 Achsen 1.3 Schwungarm 1.4 Kugelhaltung	in Stoßrichtung, parallel, schließt am Hals
	2. Stand- waage	2.1 Schulterachse 2.2 Gleitfuß 2.3 Schwungbein 2.4 Gleichgewicht	boden-parallel unverändert ausgleichend ungestört
	3. Starthocke	3.1 Beugewinkel Beine/Rumpf 3.2 Schwungarm 3.3 Schwungbeinknie	„normal“ schließt Nähe Gleitbein
	4. Angleiten	4.1 Schwungbein 4.2 Oberkörper 4.3 Schwungarm 4.4 Gleitfuß	flach/gestreckt unverändert passiv/schließt flach
	5. Stoßaus- lage	5.1 Gleitfußposition 5.2 Gleitbein 5.3 Schulterachse 5.4 Schwungfuß	eingedreht gebeugt quer am Balken

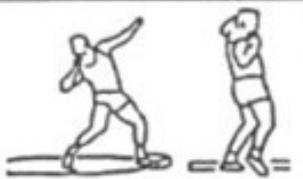
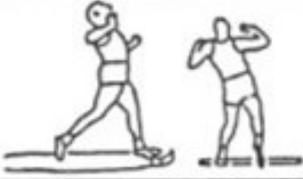
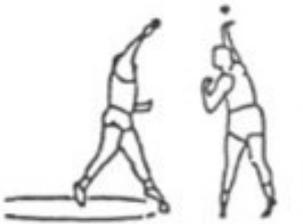
	6. Heben	6.1 Gleitfuß 6.2 Oberkörper 6.3 Schulterachse	ausdrehend aufrichtend in Stoßrichtung
	7. Ausstoß- position	7.1 Stoßarmellbogen, 7.2 Oberkörper 7.3 Gleitfußferse	angehoben Brustvorspannung ausgedreht
	8. Ausstoß	8.1 Stoßhand 8.2 Becken 8.3 Beine 8.4 linke Seite	„Ausklinken“ im Druckbereich gestreckt fixiert
	9. Dynamik	9.1 Angleiten 9.2 Stoßauslage 9.3 Ausstoß	aktiv ohne Pause explosiv

Abb. 29: Analysebogen-Grundmodell der O’Brien-Technik im Kugelstoßen (nach Tidow, 1996b, S. 257)

Da aus dem Analysebogen aber noch keine wirklichen Zusammenhänge einzelner Bewegungselemente abgeleitet werden können, beschreibt Tidow (1983) sehr genau, worauf es bei der Ausführung einzelner Phaselemente ankommt. Je besser dieses Hintergrundwissen ausgebildet ist, desto präziser wird die entsprechende Bewegungsphase mit dem Analysebogen beobachtet werden können.

Wie realistisch es erscheint, diese 32 Einzelkriterien innerhalb eines Beobachtungsversuches zu erfassen, wurde bereits im Kapitel 2.2.2 (Sportspezifische Sehstrategien) behandelt. Jedenfalls liefert das Modell eine hohe Sollwertpräzision, die als Voraussetzung für eine gezielte Bewegungsbeobachtung betrachtet werden kann. Ein sinnvoller Schritt könnte nun darin bestehen, die critical features in eine überschaubare Anzahl an Technik-Cues zu übersetzen und sich dabei auf die für das Bewegungsziel wichtigen Phaselemente zu beziehen, wie dies ist in Tab. 5 in einem Cue-basierten Analysebogen von Fronske (2014) für den Kugelstoß umgesetzt wird. Der Vorteil des letzteren Modells besteht darin, dass die vielen Einzelkriterien in weniger und zugleich leichter zu beobachtenden Cues komprimiert werden.

Tab. 5: Cue-basierter Beobachtungsbogen für die O’Brien-Technik im Kugelstoßen (nach Fronske, 2014, S. 368f.)

Skill	Cue	Why?	Common Error
Enter/Exit Position	Enter ring from same place everytime	Develop routine	Inconsistent in where entering the ring
	Exit from rear	The rule	Not exiting out the back
Hold Shot	Cradle shot	Allows for better release	Dropping the elbow and having the shot roll into palm
	Hold shot on finger pads	Best elastic release	Holding with palm
	Push shot firmly against neck	Control over shot	Shot falls out
	Elbow up	Power, control	Throwing like a baseball

Release of Shot	Push shot through head		Throwing like a baseball
	Shot leaves from neck, a ‚put‘ not a throw	Power, control	
	Release is fast like a punch and at a 45° angle	Best flight angle for maximum distance	Punch is horizontal
Start of Glide	Bend support knee	Best power	Hurrying start and keeping weight back
	Kick backward and down with nonsupport leg	Best power	Trying to throw before reaching power position
	Land both feet simultaneously still facing back	Position for greatest power	Facing sector after throw
	Hips open up quick like a baseball swing	Power	Not opening hips
	Free arm swings up as if slapping a giant	Creates torque	Not slapping high enough
Explosion of Glide	Push with both legs	Makes throw more powerful	Legs collapse on throw
	Keep shot back last	Best Power	Shot goes first
	Release shot at 45° angle	Best angle for maximum distance	Releasing too low or too high
End of Throw	Slap and pull with free arm	Shorten lever to speed rotation	Do not use off hand

2.3.2.4 Adressatenbezug

Wissen über die motorische Entwicklung sowie sensible Phasen (key milestones) befähigt die Lehrkraft außerdem dazu, in frühen Stufen des motorischen Lernprozesses die richtigen Schlüsselemente entsprechend des Entwicklungsstandes seiner Zielgruppe zu beobachten. Außerdem wird die Bewegungsbeobachtung umso effektiver, je besser ein Trainer über die häufigsten Fehler, die bei der zu unterrichtenden Bewegung auftreten können, Bescheid weiß. Im Idealfall kann er die Fehlerbilder den entsprechenden critical features zuordnen, um dem Athleten zielgerichtet verschiedenste Rückmeldungen geben zu können über seine Bewegungsausführung (Knudson, 2013, S. 85f.). Dabei sollen beobachtete Abweichungen vom altersgemäß wünschenswerten Bewegungsablauf hierarchisiert werden, d. h. eingeschätzt werden, ob und wie stark sich die Abweichung negativ auf das Bewegungsziel auswirkt. Hauptfehler, welche andere Bewegungsfehler nach sich ziehen, sind dem Sportler sofort rückzumelden und sollen so schnell wie möglich behoben werden, ehe speziellere Fehlerbilder angesprochen werden (Logan & McKinney, 1977, S. 215). Radford (1990, S. 9) empfiehlt dazu, die Bewegungsvorstellung nicht nur anhand idealer Technikdemonstrationen von Könnern zu vermitteln, sondern sich am motorischen Leistungsniveau der zu beobachtenden Zielgruppe zu orientieren, welche die Bewegungen zumeist langsamer, weniger dynamisch oder mit geringerem Bewegungsumfang ausführt. Speziell bei der Einschätzung, ob eine Bewegung noch als korrekt ausgeführt gelten kann, ist der Vergleich mit alters- oder fähigkeitskohärenten Bewegungsvorbildern realistischer als wenn Topathleten eine Bewegungsausführung nachahmen, die eigentlich unter ihrem motorischen Könnensniveau liegt. Daher wird in dieser Studie darauf geachtet, dass die Vermittlung des Sollwerts anhand von Athletinnen und Athleten erfolgt, die altersmäßig zu den Schülerinnen und Schülern passen, deren Wurfbewegung im Beurteilungstest bewertet werden muss, diesen aber aufgrund ihrer leichtathletischen Grundbildung vom technischen Niveau her deutlich überlegen sind und als Bewegungsvorbild für Gleichaltrige dienen können.

2.3.2.5 Zusammenfassung

Es lässt sich zusammenfassen, dass eine möglichst präzise Sollwertvorstellung eine Voraussetzung zu sein scheint für eine gute Bewegungsbeurteilung. Dabei geht es aber nicht nur um die Kenntnis von räumlich-zeitlichen Abläufen, sondern auch um ein tieferes biomechanisches Verständnis, um einzelne Schlüsselemente in ihrer Wichtigkeit für die Gesamtbewegung richtig einzuschätzen und eine für die Adressatengruppe passende Hierarchisierung vorzunehmen. Dabei kann es helfen, häufig vorkommende Abweichungen vom idealen Bewegungsverlauf zu kennen, aber auch angemessen einzuschätzen, wie stark sich die Bewegungsabweichung auf das Erreichen des Bewegungsziels auswirkt. Für die Konstruktion des Fachwissenstests bedeutet das, dass neben dem räumlich-zeitlichen Wissen über die Bewegung (siehe Kapitel 4.2.1) auch biomechanische Inhalte und Zusammenhänge (siehe Kapitel 4.2.2) sowie Kenntnisse über Fehlerbilder (siehe Kapitel 4.2.3) abgefragt werden, um die Sollwertpräzision in möglichst vielen Facetten zu erfassen und mit der Beurteilungsleistung in Beziehung zu setzen. Das Wiedererkennen von Schlüsselpositionen sowie deren korrekte Abfolge wird in enger Anlehnung an die Verwendung von Piktogrammen in Technikanalysebögen mittels eines Bildkartentests überprüft und soll die visuelle Bewegungsvorstellung operationalisieren (siehe Kapitel 4.2.4).

Bei der Durchführung des Beurteilungstests soll darauf geachtet werden, dass den Testpersonen der Sollwert der Schlagwurfbewegung altersgemäß vermittelt wird, indem Bewegungsdemonstrationen von Athletinnen und Athleten verwendet werden, die den Schlagwurf für ihr Alter schon sehr gut beherrschen. Auch soll neben der idealen Bewegungsausführung anhand von Videobeispielen konkretisiert werden, welche technischen Fehl Ausführungen auf eine geringe oder starke Abweichung vom Sollwert hinweisen (siehe die Videos zur Erklärung der Bewegungsmerkmale als Einstieg in den Beurteilungstest). Damit soll sichergestellt werden, dass die Testpersonen eine möglichst einheitliche Bewegungsvorstellung vom idealen, altersgemäßen Schlagwurf ausbilden und die beiden Abweichungsstufen in gleicher Weise einschätzen. Die Gruppierung der Technikdetails zu verschiedenen Bewegungsmerkmalen (nach dem Vorbild Meinels in Harmonie, Rhythmus, Tempo, Umfang, Kopplung und Präzision) soll den Beurteilenden helfen, trotz der Beobachtung einzelner Bewegungsmerkmale die Gesamtbewegung nicht aus dem Blick zu verlieren.

Die Fähigkeit, das Bewegungswissen wieder in einfacher zu verstehende Bewegungs-Cues umzusetzen, wird in dieser Studie nicht erfasst, weil keine unmittelbare Auswirkung auf die Bewegungsbeurteilung angenommen wird. Wenn es aber darum geht, aus der Bewegungsbeurteilung Instruktionen und Feedback für den zu beobachtenden Sportler abzuleiten, dann ist diese Fähigkeit zur didaktischen Reduktion von entscheidender Bedeutung.

2.3.3 Bewegungserfahrung (Eigenrealisation)

Wird unter Bewegungserfahrung der „durch motorische Lern- und Übungsprozesse angeeignete Gesamtbestand von Bewegungsmustern und damit zusammenhängende Einsichten“ (Fetz, 1980, S. 208) verstanden, so kann daraus abgeleitet werden, dass sich neben den sensomotorischen Elementen der Eigenrealisation gerade ihre kognitiven Anteile auf die Verfeinerung der Bewegungsvorstellung auswirken können. Neumaier (1988, S. 286) weist darauf hin, dass ein hohes Maß an Bewegungserfahrung zwar eine notwendige, aber keine hinreichende Voraussetzung für das Zustandekommen eines hohen Eigenrealisationsniveaus ist, weshalb beide Begriffe voneinander unterschieden werden müssten. Bewegungserfahrung ist für ihn daher zu verstehen als „Summe der praktischen, aktiven Beschäftigungszeit mit einer bestimmten sportlichen Technik“.

Neumaier (1988, S. 284f.) betont, dass „die Vorstellung von einer sportlichen Technik [...] ‚fragmentarisch‘ bleibt, wenn sie ausschließlich eigenrealisationsfrei erworbene Informationen umfaßt [sic!].“ Die Qualität der Bewegungsbeurteilung leidet, wenn keine motorische Identifikation mit der zu beurteilenden Bewegung vorliegt, und der vorzunehmende Soll-Istwertvergleich wird vor allem in dynamischer Hinsicht unvollkommen, weil die Komponenten Richtung, Größe, Zeitpunkt oder Dauer der Impulsgebung nicht korrekt eingeschätzt werden können. Gerade diese Fähigkeit des Mitvollziehens einer Bewegung bestimme zu einem hohen Anteil die Trefferquote bei der Analyse der Fehlerursache (Tidow, 1981, S. 264). Trotzdem kann eine intensive, ausschließlich theoretische Schulung der Bewegungsvorstellung eine deutliche Verbesserung der Beurteilungsleistung zur Folge haben, auch wenn nicht geklärt ist, bis zu welchem Beurteilungsniveau auf diese Weise eine in Verbindung mit der Eigenrealisation gebildete Bewegungsvorstellung ausgeglichen werden kann. Die Verbindung beider Komponenten lässt jedenfalls die günstigsten Ergebnisse erwarten (Neumaier, 1988, S. 297).

Zudem bewirkt ein gutes Eigenrealisationsniveau, dass eine Bewegung in ihren ursächlichen Zusammenhängen und Wirkungen besser verstanden und gedanklich nachvollzogen werden kann (Pöhlmann, 1986, S. 120). Das Herausfinden von Fehlerursachen setzt wiederum das Erkennen von Abweichungen in der sichtbaren Bewegungsstruktur voraus. Eine derart anspruchsvolle Beurteilungsaufgabe (wie das Herausfinden von Hauptfehlern, welche die weitere Bewegungsausführung negativ beeinflussen) kann aber selbst für Beurteilende, die den Bewegungsablauf sicher beherrschen, eine Überforderung darstellen (Neumaier, 1988, S. 296).

Ob das Eigenrealisationsniveau die Fähigkeit zur Bewegungsbeurteilung maßgeblich bestimmt, ist in der Forschung durchaus umstritten. Schon 1967 stellen sich Girardin und Hanson (1967, S. 556) die Frage: „How important is personal physical skill to the effective teaching of physical activity skill?“ Sie untersuchten die Beziehung zwischen der Bewegungsdemonstration von elf Bodenturnelementen und der Fähigkeit, Ausführungsfehler derselben diagnostizieren zu können. Sie kamen zu dem Ergebnis, dass eine Korrelation besteht zwischen der Eigenrealisation und der Beurteilungsfähigkeit sowie zwischen der theoretischen Kenntnis über die Bewegung und der Beurteilungsfähigkeit. Osborne und Gordon (1972) hingegen konnten bei 90 Versuchspersonen keinen Einfluss der Ausführungsqualität des Vorhandschlags im Tennis auf die Beurteilungsqualität nachweisen. Neumaier (1988, S. 294–296) fasst seine Untersuchungsergebnisse so zusammen, dass die mit der Verbesserung der Bewegungsqualität zunehmende Differenzierung der bewusst verfügbaren Bewegungsvorstellungen (z. B. Wissen bezüglich der Bedeutung von Ausführungsmerkmalen zu den verschiedenen Bewegungsparametern) sich positiv auf die Beurteilungsfähigkeit auswirkt. Die Beurteilungsfähigkeit steigt dabei aber nicht stetig

mit zunehmendem motorischen Können an, sondern erreicht gelegentlich schon auf geringem technischem Ausführungsniveau ein Leistungsplateau (z. B. beim Sprungwurf im Basketball). Je komplexer die Bewegung jedoch auszuführen bzw. je schwieriger sie zu erlernen ist (z. B. Stabhochsprung), desto höher muss das motorische Können ausgeprägt sein, damit die Bewegungsvorstellung präzise genug ist, um die Beurteilungsleistung positiv beeinflussen zu können.

Möglicherweise kann sich das Eigenrealisationsniveau indirekt positiv auf die Beurteilungsleistung auswirken, wenn ein positiver Zusammenhang zum detaillierten Wissen um Bewegungsmerkmale im Sinne einer bewussten oder bewusstseinsfähigen Bewegungsvorstellung nachgewiesen werden kann. Girardin und Hanson (1967, S. 559) fanden in oben genannter Untersuchung keinen Zusammenhang zwischen der Eigenrealisation und dem theoretischen Bewegungswissen. Neumaier (1988) hingegen nimmt an, dass sich „ein gutes Eigenrealisationsniveau in der überwiegenden Zahl der Fälle auch in einem differenzierteren Wissen um die Merkmale der Idealbewegung ausdrückt“ (S. 288) und somit bei guten Sportlern auch bessere Voraussetzungen für eine hohe Beurteilungsqualität vorliegen.

Zu widersprüchlichen Ergebnissen kommen Neumaier (1979, S. 202) und Mester (1988, S. 176) bei ihren Untersuchungen bezüglich der Beobachtung von Übungsteilen aus dem Gerätturnen bzw. Tennis. Während ersterer feststellt, dass Leistungsturner bei der Bewegungsbeobachtung signifikant mehr antizipatorische Sakkaden ausführen als Nichtturner, widerspricht zweiterer, dass es bei Tennisspielern mit zunehmendem Eigenrealisationsniveau als sinnesphysiologische Anpassungserscheinungen während des Beurteilungsvorgangs zu einer **Reduzierung der Anzahl sakkadischer Augenbewegungen** kommt. Neumaier (1988, S. 298) bemerkt in einer späteren Untersuchung dazu folgendes:

Es kommt demnach mit steigendem Könnensniveau in vielen Fällen zu einer ‚Beruhigung‘ des Blickverhaltens, die als eine Ökonomisierung der Informationsaufnahme zu interpretieren ist. Blickbewegungen werden bei Personen mit einem geringen sportartspezifischen Könnensniveau, d. h. insbesondere einer mangelhaften Bewegungsvorstellung, stärker von zufallsbedingten Faktoren und Einflüssen [...] gesteuert [...]. Dadurch werden häufig mehr Sakkaden als sinnvoll ausgelöst, was aufgrund der sakkadischen Suppression zu Wahrnehmungsbeeinträchtigungen bei gleichzeitig verringerter Fixationsdauer auf den für die Aufgabenlösung informationsträchtigen Elementen führt.

Diese Anpassungserscheinung kann bereits (zumindest bei einfach zu erlernenden Fertigkeiten) in frühen Phasen des motorischen Lernprozesses auf noch relativ geringem Könnensniveau ebenso nachgewiesen werden wie nach einer eigenrealisationsfreien Präzisierung der Bewegungsvorstellung.

Bezüglich der **Wahl der Fixationsorte** während der Beurteilungsaufgabe kann keine eindeutige Abhängigkeit vom Eigenrealisationsniveau konstatiert werden, denn die Fixationsorte streuen unabhängig vom sportlichen Können der untersuchten Beobachter sportartübergreifend relativ stark über den gesamten Körper. Zwar liegt es nahe, dass Bewegungsexperten eher als Anfänger wissen, worauf sie bei der Beurteilung besonders achten sollen, jedoch drückt sich diese Vermutung nicht in interindividuell übereinstimmenden Fixationsbereichen für die Bewältigung gestellter Beurteilungsaufgaben aus. Dies führt Neumaier (1988, S. 299) auf verschiedene Gründe zurück:

- Die Beobachter konzentrieren sich auf verschiedene Merkmale der Bewegung, die nicht alle gleichzeitig beobachtet werden können. Je nach subjektiver Gewichtung dieser Merkmale ergeben sich auch unterschiedliche Fixationsziele.
- Die Blicksteuerung geschieht überwiegend unbewusst, wenn dynamische Aspekte wie Bewegungsfluss, Bewegungsrhythmus oder Timing von Bewegungskopplungen beurteilt werden sollen. Auch wenn das gleiche Merkmal beobachtet werden soll, richten Beobachter deshalb nicht automatisch die Fixation auf die gleichen Körperregionen.

- Der Sehwinkel des Beobachters, d. h. die Ausdehnung des Objektdetails (Sportler, Körperteil) im Gesichtsfeld, bestimmt maßgeblich die Wahl der Fixationsorte. Der Beobachter kann unnötige Blickbewegungen bei synchronoptischen Identifikationsaufgaben vermeiden, wenn das Objekt größtenteils foveal bzw. parafoveal (Abweichungen im Bereich von 2–5°) abgebildet wird.

Bei der Bewegungsbeobachtung zum Zwecke der Fehleranalyse neigen Beurteiler häufig dazu, einen Fixationsbereich zu wählen, bei dem die Wahrscheinlichkeit des Übersehens wichtiger Fehler am geringsten ist, d. h. bei dem potentiell fehlerträchtige Regionen über die para-foveale Informationsaufnahme noch zufriedenstellend identifizierbar sind. Die Blickstrategie hängt dabei stark von der subjektiv eingeschätzten Funktion und Wichtigkeit einzelner Aktionen und Positionen im Bewegungsablauf ab. Auch die zur Verfügung stehende Beobachtungszeit kann bei schnell ablaufenden Bewegungen so gering sein, dass Fixationswechsel weder sinnvoll noch möglich sind und ein möglichst zentraler Fixationspunkt eingenommen wird. Wenn deutliche Fehler bei schnell ablaufenden azyklischen Bewegungen auftreten, die dem Beurteiler vorher nicht bekannt sind, beeinflussen diese den Beobachtungsvorgang nicht. Es kommt also nicht zu Blicksprüngen auf Körperregionen, in denen ein Fehler sichtbar lokalisiert ist, denn dies würde die weitere Informationsaufnahme behindern sowie möglicherweise sogar zu einer retroaktiven Löschung kurz zuvor aufgenommener Informationen über den Bewegungsfehler führen (Neumaier, 1988, S. 298–301).

Die Eigenrealisation scheint demnach ein wichtiger Einflussfaktor auf die Beurteilungsleistung zu sein, wenngleich über die konkrete Bedeutung auf den Beurteilungsprozess wenig bekannt ist. Daher soll die Qualität der Eigenrealisation der beurteilenden Testpersonen über ein videogestütztes Analyseverfahren gemessen werden (siehe Kapitel 4.4.1), um zu überprüfen, ob ein Zusammenhang mit der Fremdbeurteilung von Schlagwurfbewegungen festzustellen ist. Dazu kommt der schon in Kapitel 2.3.2.3 erwähnte Detailanalysebogen zum Einsatz, in welchem 39 Technikkriterien eingeschätzt oder gemessen werden, um die Qualität der Eigenrealisation möglichst objektiv feststellen zu können.

2.4 Schlussfolgerungen für die methodische Konzeption dieser Studie

Am Ende der theoretischen Beschäftigung mit den Begriffen der Bewegungsbeobachtung und Bewegungsbeurteilung soll nochmals zusammengefasst werden, wie die gewonnen Erkenntnisse für den weiteren Verlauf dieser Studie genutzt werden können. In Kapitel 2.1 wurde die Bewegungsbeurteilung in den größeren Zusammenhang der fachdidaktischen Diagnosekompetenz eingeordnet, um den situativen Kontext zu klären, in welchem die Beobachtungs- und Beurteilungskompetenz von Sportlehrkräften sowie Trainerinnen und Trainern täglich angewendet werden muss. Kernelement stellt dabei die morphologische Betrachtungsweise dar, bei der sportliche Bewegungen ohne technische Hilfsmittel lediglich mit dem optischen und akustischen Analysator beobachtet werden. Den Schwierigkeiten diagnostischer Prozesse im Sport wie die meist sehr kurze, nicht wiederholbare Bewegungspräsentation oder die Wahrnehmung zum Teil synchron ablaufender Körperteilbewegungen muss ein Beobachter angemessen begegnen können, um zu einer korrekten Beurteilung des Bewegungsablaufs zu kommen. Umso wichtiger ist es, den Beurteilenden bei der Messung der Beurteilungsleistung nicht zu überfordern, indem sowohl einfach nachvollziehbare Beurteilungskriterien ausgewählt als auch bewältigbare Beurteilungsaufgaben geschaffen werden. Daher werden von einem hierarchischen Grundmodell allgemeiner und spezifischer Struktur- und Güteaspekte sechs Meinel'sche Bewegungsmerkmale ausgewählt, anhand derer die Beurteilung der leichtathletischen Schlagwurfbewegung erfolgen soll (siehe Kapitel 2.1.3.3.5 und 4.5.1): die Bewegungsharmonie, das Bewegungstempo, der Bewegungsrhythmus, der Bewegungsumfang, die Bewegungskopplung und die Bewegungspräzision.

Die Beurteilungsleistung lässt sich dabei in zwei Kompetenzbereiche unterteilen: das Beobachten und das Beurteilen. Hinsichtlich der Beobachtungskompetenz sind vor allem die sinnesphysiologischen Limitationen zu beachten, welche den optischen Analysator bei sportlichen Bewegungen in seiner Aufnahmefähigkeit begrenzen. Das Auge kann nur optische Reize, die innerhalb der Fovea centralis abgebildet werden, mit bestmöglicher Sehschärfe wahrnehmen. Da es sich beim leichtathletischen Schlagwurf um eine dynamische, in gewissen Phasen sehr schnell ablaufende Bewegung handelt, ist der Beobachter gefordert, den Werfer (je nach Entfernung zum Beobachtungsobjekt) mittels Blick-, Augen- und Kopfbewegungen so zu verfolgen, dass der Bereich des schärfsten Sehens auf die Körperregionen gerichtet ist, an denen sich die entscheidenden Bewegungen einer bestimmten Phase abspielen (dynamische Sehschärfe, siehe Kapitel 2.2.1). Man spricht bei solchen Blicknachfolgebewegungen auch vom efferenten Bewegungssehen. Dabei kann nicht der ganze Körper scharf betrachtet werden, sondern der Beobachter muss sich entscheiden, wohin er seine Blickfixation richtet, weil der hundertprozentige Schärfbereich nur einen sehr kleinen Winkelbereich von 1° umfasst. Alle außerhalb der Fovea centralis liegenden Reize können somit nur mit einem sehr starken Sehschärfeverlust verarbeitet werden, was natürlich zu Informationsverlusten und Fehleinschätzungen dieser lediglich peripher wahrgenommenen Bewegungsausführungen führen kann. Diese Einschränkung betrifft vor allem die synchroptische Bewegungswahrnehmung (siehe Kapitel 2.2.1.2), wenn es darum geht, gleichzeitig ablaufende Teilbewegungen zu beobachten und zu beurteilen, wie dies beispielsweise bei der Ausführung des Impulsschrittes mit gleichzeitiger Fokussierung auf die Wurfarmstreckung der Fall wäre. Um diesen Wahrnehmungskonflikt zu lösen, wenden erfahrene Beobachter spezielle Sehstrategien an (siehe Kapitel 2.2.2.2). Jedoch sind die Erkenntnisse über die angewandten Sehstrategien nicht so eindeutig, dass sich daraus ein nennenswerter Einfluss auf die Beurteilungsleistung ableiten ließe, weshalb in dieser Studie die Aufzeichnung von Blickbewegungen mittels Eyetracking-Systemen nicht zum Einsatz kommt. Es kann also schlussgefolgert werden, dass die „Optimierung des Blickverhaltens, insbesondere das Vermeiden von zu vielen, unnötigen Blickbewegungen [...] eine notwendige, wenn auch

nicht hinreichende Voraussetzung für die optimale visuelle Informationsaufnahme und -verarbeitung im Beurteilungsprozeß [sic!]“ (Neumaier, 1988, S. 325) ist, welcher „nur in Verbindung mit einer aufgabenadäquaten Sollwertbildung“ (Neumaier, 1988, S. 378) zufriedenstellend zu erfüllen ist.

Da die Einzelphasen des Schlagwurfs sehr schnell aufeinander folgen und mitunter innerhalb kürzester Zeit die Körperregion gewechselt werden muss (z. B. vom Fußaufsatz des Stemmbeins zum Wurfmellbogen), kommt es zu sog. Sakkaden, bei denen der Blick zwar sehr schnell von einem Beobachtungspunkt zu einem anderen gewechselt werden kann, jedoch in dieser Zeit keine Informationsaufnahme erfolgen kann. Pro Sekunde können nur 2–4 Sakkaden ausgelöst werden, so dass ein zu häufiger Wechsel des Fixationspunkts nicht empfehlenswert ist. Erfahrene Beobachter nutzen Sakkaden vor allem dazu, einen zwischenzeitlichen Verlust der Blickfixierung zu kompensieren und das Auge antizipatorisch auf den nächsten Beobachtungspunkt zu lenken, um dort die Informationsaufnahme wieder rechtzeitig starten zu können. Außerdem bewirkt eine gut ausgebildete Analogiefähigkeit, dass von einer beobachteten Teilbewegung auf die Ausführung unbeobachteter Bewegungsmerkmale geschlossen werden kann, wodurch sowohl die Quantität als auch die Qualität der Beurteilungsleistung zunimmt. Da die zu beurteilenden Bewegungen bei der Messung der Beurteilungsleistung mittels Videos auf einem großen Smart-TV präsentiert werden, gilt es im Vergleich zur Realbeobachtung zu bedenken, dass bei der Videobeobachtung durch nicht notwendige Kopfbewegungen die Blickfixierung vereinfacht ist, weil die Augen schneller bewegt werden können als der Kopf bzw. der Körper. Aufgrund der Größe der Projektionsfläche werden aber trotzdem Sakkaden notwendig sein, um während der gesamten Schlagwurfbewegung die entscheidenden Körperteile möglichst scharf wahrnehmen zu können. Diese spezielle Beobachtungssituation gilt es bei der späteren Einordnung der Ergebnisse bezüglich der Messung der Beurteilungsleistung zu bedenken.

Es stellt sich abschließend die Frage, von welchen möglichen Faktoren die Beurteilungsleistung generell abhängt: Wie wichtig ist dabei die Erfahrung in der Beurteilung von Bewegungsabläufen im Vergleich zu Beobachtern ohne spezifische Lehrerfahrungen? Welchen Anteil nehmen dabei Erfahrungen aus der Eigenrealisation in der betreffenden Sportart ein? Und gibt es einen sog. Beobachtungstransfer, d. h. inwieweit profitiert man von sportartspezifischen Beobachtungserfahrungen bei der Beurteilung ungewohnter sportlicher Bewegungen. Die Forschungslage dazu ist recht spärlich und es gibt wenige aktuelle Untersuchungen.

Neumaier (1988, S. 318–325) fasst die Untersuchungsergebnisse, die im Rahmen einiger Diplomarbeiten an der Deutschen Sporthochschule Köln in den 1980er Jahren durchgeführt wurden, in folgenden Punkten zusammen:

- „Trainer bzw. Experten sind solchen Personen, die weder eine sportartspezifische Lehrtätigkeit ausüben noch ein hohes Eigenrealisationsniveau in den betreffenden sportlichen Techniken besitzen, in der Bewegungsbeurteilung überlegen.“ (Neumaier, 1988, S. 318)
- Trainer haben gegenüber fortgeschrittenen Sportlern mit hohem Eigenrealisationsniveau keine grundsätzlichen Vorteile bei der Bewegungsbeurteilung. Beide Gruppierungen wissen, worauf es bei der Bewegungsausführung ankommt und können dies auch dementsprechend beobachten.³²
- Die Zeitdauer (in Jahren) der Lehr- oder Beurteilungstätigkeit allein liefert keinen Erklärungswert für das Könnensniveau eines Beurteilers. Vielmehr bestimmt das Klientel, mit dem ein Trainer täglich arbeitet,

³² Neumaier (1988, S.303f.) verweist auf einen unveröffentlichten Forschungsbericht von Daus, Blischke, Olivier und Graf (1983), wonach erfahrene Leichtathletiktrainer im Vergleich zu aktiven Kugelstoßern die gleiche Beurteilungsleistung beim Bewerten von Kugelstoßbewegungen erzielten, die Trainer jedoch ihren Blick häufiger auf den Bereich der unteren Extremitäten richteten (61,6 %) als die Athleten, welche die oberen und unteren Körperregionen gleich oft fixierten.

- maßgeblich die angewendeten Sehstrategien sowie die Beurteilungsleistung. Trainer, die häufig Athleten höherer Leistungsstufen beobachten, achten auf andere Bewegungsdetails als Trainer von Athleten niedrigerer Leistungsstufen. Dies kann auch zu unerwünschter Habitualisierung des Blickverhaltens führen, das dann bei fremden Beurteilungsobjekten möglicherweise unangemessen ist (z. B. durch Beobachtung von Bewegungsdetails, die auf einer bestimmten Könnensstufe keine Relevanz mehr haben)
- Ob spezifische Lehr- und Beurteilungstätigkeiten sich positiv auf die Fähigkeit zur Bewegungsbeurteilung auswirken, hängt davon ab, wie gut bei einer Lehrkraft die kognitive Strukturierung und Differenzierung einer Bewegung mit einer hieraus entstehenden detaillierten, bewusst verfügbaren Bewegungsvorstellung ausgebildet ist.
 - Bei der Beurteilung komplexer Bewegungsabläufe gibt es keine allgemeine Überlegenheit von Personen, die über Erfahrungen aus Beurteilungstätigkeiten in Verbindung mit anderen Bewegungsabläufen stammen. Entscheidend für das Erkennen von spezifischen Bewegungsfehlern ist das Vorhandensein eines entsprechenden Sollwertes beim Beurteiler sowie seiner Fähigkeit zum inneren Ist-Sollwert-Vergleich.
 - Die Wahrnehmung kinematischer Merkmale (z. B. Anlaufgeschwindigkeit) hat eine hohe Aufgabenschwierigkeit und kann von Novizen genauso gut bzw. schlecht beobachtet werden wie von erfahrenen Trainern. Somit liegt der Schluss nahe, dass diese Einschätzung auf einer allgemeineren Beobachtungsfähigkeit oder gar mehreren bewegungsspezifischen Beobachtungsfähigkeiten beruht.
 - Bezüglich der Fixationspräferenzen in Abhängigkeit von der Beurteilungserfahrung bzw. dem Qualifikationsniveau der Beobachter gibt es kein einheitliches Gesamtbild. Unterschiedliche Fixationspräferenzen bei gleichen Beurteilungsleistungen sind ebenso nachgewiesen worden wie gleiche Fixationspräferenzen bei unterschiedlichen Beurteilungsleistungen. Auch für die Tatsache, dass unterschiedliche Beurteilungen auf verschiedenen Schwerpunkten bei den Fixationen zurückzuführen sind oder gleiche Beurteilungsleistungen mit ähnlichen Fixationsstrategien erzielt werden, gibt es Belege. Somit liegt die Vermutung nahe, dass der innere Vergleich der wahrgenommenen Bewegung mit dem Sollwert zur richtigen Bewegungsausführung für den Beurteilungsvorgang entscheidend ist.
 - Defizite im Sollwert verhindern also die adäquate Bewertung (oder sogar schon Aufnahme) von Informationen über den Bewegungsablauf
 - „Mit der Verbesserung der Bewegungsvorstellung kommt es relativ früh zu einer individuellen Ausprägung von Präferenzen für die Bewegungsteile, auf die man besonders achtet. Dies zieht verlängerte Fixationszeiten, d. h. auch eine Reduzierung der Sakkadenzahl pro Zeiteinheit nach sich.“ (Neumaier, 1988, S. 325)
 - Schon ein relativ niedriges Niveau der Eigenrealisation (z. B. nach wenigen Unterrichtseinheiten zum Erlernen einer Bewegung) reicht aus, um eine signifikante Senkung der Sakkadenanzahl in der entsprechenden Beurteilungsaufgabe zu erzielen.

Die Beurteilungserfahrung erweist sich demnach als möglicher Einflussfaktor auf die Beurteilungsleistung. Dabei ist zu vermuten, dass es eine Rolle spielt, ob die Beurteilungserfahrung direkt anhand der leichtathletischen Schlagwurfbewegung oder lediglich anhand schlagwurfähnlicher Bewegungen aus anderen Sportarten (z. B. Schmetterschlag im Volleyball) gesammelt wurde. Besonders gering oder fast nicht vorhanden scheint der Einfluss zu sein, wenn die Beurteilungserfahrung aus Sportarten stammt, in denen überhaupt keine schlagwurfähnlichen Bewegungen vorkommen. Außerdem bleibt zu berücksichtigen, dass eine stark positive Wirkung zu erwarten ist, wenn die Beurteilungserfahrung an Personen erworben wurde, die den zu beurteilenden Sportlerinnen und Sportlern (beispielsweise bezüglich Alter oder Leistungsniveau) dieser Studie ähneln. Auch das Eigenrealisationsniveau scheint die Beurteilungsleistung zu beeinflussen: bereits auf einem geringen Könnensniveau kommt es auch zu einer effektiveren Sehstrategie mit weniger Sakkaden, ein hohes Bewegungskönnen wirkt sich sowohl bei Trainerinnen und Trainern als auch bei Athletinnen und Athleten positiv auf die Bewegungsbeurteilung

aus. Eine besonders wichtige Rolle bei der Bewegungsbeurteilung scheint die Qualität der Bewegungsvorstellung (Sollwertpräzision) einzunehmen. Je besser ein Beurteiler über eine Bewegung, ihre Knotenpunkte und ihre Phasenstruktur Bescheid weiß oder sogar biomechanische Zusammenhänge kennt, desto besser ist die Beurteilungsqualität oder desto zuverlässiger können Bewegungsabweichungen erkannt werden. Auch die Fähigkeit zum inneren Ist-Sollwert-Vergleich kann die Bewegungsbeurteilung beeinflussen, wenn es darum geht, eine real präsentierte Bewegung mit einer Bewegungsvorstellung zu vergleichen. Diese allgemeine Fähigkeit kann beispielsweise mittels des Tests zur Kontrolle der Bewegungsvorstellungsfähigkeit (TKBV) gemessen werden.

Somit entstammen aus der theoretischen Auseinandersetzung mit der Bewegungsbeobachtung und Bewegungsbeurteilung schon einige mögliche Einflussfaktoren, die zur gemessenen Beurteilungsleistung in Beziehung gesetzt werden können: das Wissen über die Bewegung des Schlagwurfs, die Qualität der Eigenrealisation, die Fähigkeit zum Ist-Sollwert-Vergleich sowie die Beschaffenheit der Beurteilungserfahrung.

Ein letzter Faktor, der die Bewegungsbeurteilung positiv beeinflusst, ist die Fähigkeit, sich trotz verschiedener Störungen (z. B. anderer Bewegungen) auf die Beobachtungsaufgabe konzentrieren zu können (Gangstead & Beveridge, 1984, S. 60). Barrett (1979) berichtet, dass Studierende bei der Reflexion über die eigene Beobachtungsfähigkeit oft berichten, wie leicht sie sich beim Beobachten abgelenkt fühlen. Ursachen dafür können verschiedenartig sein:

Knowledge of movement, your role in the setting, fear, excitement, worry, type/amount of equipment/apparatus, number of students, size of teaching area; skill level of students; and speed, complexity, and number of times a movement is seen. (Barrett, 1984, S. 29)

Insofern finden Lehrkräfte im Sportunterricht immer ungünstige und teilweise unkontrollierbare Beobachtungsbedingungen vor, weil die Aufmerksamkeit meist nicht allein auf die Bewegungsbeobachtung gerichtet werden kann, sondern parallel dazu noch andere Herausforderungen gemeistert werden müssen. Die in dieser Studie verwendete Leistungsmessung mit Videovignetten reduziert die äußeren Störgrößen, die bei einer realen Bewegungsbeobachtung auftreten können, da die Bewegungsbeurteilung sozusagen unter Laborbedingungen in einem Multimedia-Raum stattfindet, weshalb dieser Faktor unberücksichtigt bleiben kann. Trotzdem entfernt sich die Messung der Beurteilungsleistung durch ihre Zweidimensionalität von der eigentlichen Beobachtungsrealität. Es wird davon ausgegangen, dass die Beurteilungsleistung von Videovignetten auf die Realsituation übertragbar ist, ohne dass dies jedoch eine wissenschaftlich abgesicherte Annahme darstellt. Insofern muss für die Ergebnisse dieser Studie immer berücksichtigt bleiben, dass diese lediglich für die Beurteilung gefilmter Bewegungen zutreffen, nicht jedoch für eine Realbeurteilung in gleicher Weise angenommen werden dürfen.

3 Fragestellung und Hypothesenbildung

Sehr viele pädagogisch-didaktische Handlungen von Sportlehrkräften beruhen auf der Fähigkeit, sportliche Bewegungen beobachten und beurteilen zu können. Wenn Heranwachsende Bewegungen erlernen wollen, sind sie auf die korrekte Rückmeldung über ihre Bewegungsausführung angewiesen. Speziell im Sportunterricht bewerten Lehrkräfte die Bewegungsqualität ihrer Schülerinnen und Schüler häufig in Form von Schulnoten, um die Gleichwertigkeit des Faches Sport im Vergleich zu anderen Schulfächern zu legitimieren. So heißt es beispielsweise im Fachprofil Sport der bayerischen Schullehrpläne, dass im Rahmen der prozessbezogenen Kompetenz Leisten den Schülerinnen und Schülern die Fähigkeit vermittelt werden soll, „sportliches Handeln [...] in Bezug zu subjektiven und objektiven Gütekriterien zu setzen und entsprechend zu beurteilen“ (Staatsinstitut für Schulqualität und Bildungsforschung München [ISB]). Allein diese wenigen Situationen reichen um zu erkennen, dass das Beobachten und Beurteilen von Bewegungen eine Kernkompetenz des Sportlehrerhandelns darstellt. So wird in Kapitel 2.1.3 die Bewegungsbeurteilung als fachdidaktische Diagnosekompetenz dargestellt, wobei das Beobachten, Analysieren und Beurteilen nur der Anfangspunkt einer umfassenderen Kompetenz ist, die das Generieren von Rückmeldeinformationen für den Sportler ebenso umfasst wie die Notengebung oder das Inszenieren von Übungsinhalten, die zur individuellen Verbesserung des Bewegungsablaufs beitragen. Umso überraschender erscheint es, dass der Begriff Beurteilen in den Kerncurricula der inhaltlichen Prüfungsanforderungen der Lehramtsprüfungsordnung I für das Fach Sport kein einziges mal auftaucht – die Beschäftigung mit diagnostischen Prozessen wird im Bezug auf das Bewegungssehen und Bewegungslernen also in der universitären Lehrerbildung nicht explizit aufgetragen (Ausgestaltung der inhaltlichen Prüfungsanforderungen für die Erste Staatsprüfung nach Kapitel II der Lehramtsprüfungsordnung I zu den einzelnen Fächern (Kerncurricula), 2015). In der zweiten Phase der Sportlehrerausbildung wird jedoch von den Referendarinnen und Referendaren im Sportunterricht verlangt, gezielte Rückmeldungen über Bewegungsausführungen an die Schülerinnen und Schüler zu geben oder diese sogar bezüglich ihrer Bewegungsqualität zu benoten, ohne dass jene im Verlauf ihrer Ausbildung angemessen auf diese schwierige Aufgabe vorbereitet wurden.

Es kann demnach eine gewisse Kluft zwischen dem Anspruch einer fachdidaktisch reflektierten Bewegungsbeurteilung und der Wirklichkeit einer unvollkommenen diagnostischen Kompetenzentwicklung am Ende der universitären Sportlehrerbildung behauptet werden. Daraus resultiert auch die grundsätzliche Forschungsfrage, wie die Fähigkeit zur Beurteilung sportlicher Bewegungen in angemessener Weise vermittelt werden kann. In diesem Bereich herrscht ein deutliches Forschungsdefizit, da sich die Sportwissenschaft seit den Studien von Neumaier (1988) nicht mehr ernsthaft mit den Bereichen Bewegungsbeobachtung und Bewegungsbeurteilung beschäftigt hat. Daraus leitet sich auch das wissenschaftliche Interesse dieser Studie ab, sich dem Phänomen der Bewegungsbeurteilung empirisch zu nähern. Da bislang noch keine erprobten Testverfahren entwickelt wurden, um die Beurteilungsleistung messbar zu machen, soll in dieser Arbeit der Versuch unternommen werden, am Beispiel der leichtathletischen Schlagwurfbewegung verschiedene Beurteilungsaufgaben auf ihre wissenschaftliche Tauglichkeit hin zu untersuchen. Ein empirischer Zugang zur Messung der Beurteilungsleistung ist grundsätzlich möglich, wenn es gelingt, morphologisch wahrnehmbare Beurteilungsmerkmale zu begründen, die je nach Übereinstimmung oder Abweichung mit einem Expertenurteil quantifizierbar sind (siehe Kapitel 4.5). Eine Komplexitätsreduzierung der Bewegungsbeurteilung erscheint dahingehend zulässig und methodisch vertretbar, dass die Messung der Beurteilungsleistung anhand von fünf konkreten, am Unterrichtsalltag einer Sportlehrkraft orientierten Beurteilungssituationen vorgenommen

wird (siehe Kapitel 4.5.3). Anhand der Einschätzung der Bewegungsmerkmale in drei Qualitätsstufen (siehe Kapitel 4.5.2) wird lediglich das Ergebnis des Beurteilungsprozesses erfasst, nicht jedoch die bei der Bewegungsbeurteilung ablaufenden sinnesphysiologischen oder psychologischen Prozesse.

Um weiterführende Überlegungen anstellen zu können, wie die Beurteilungskompetenz geschult und verbessert werden kann, erscheint es hilfreich, sich mit möglichen Einflussfaktoren auf die Beurteilungsleistung auseinanderzusetzen und beispielsweise zu wissen, ob ein gutes Fachwissen über die zu beurteilende Bewegung oder die Qualität der Eigenrealisation die Beurteilungsleistung positiv oder negativ beeinflussen. Der praktische Nutzen dieser Erkenntnisse liegt auf der Hand: Wenn klar ist, welche Voraussetzungen erfüllt sein müssen, damit eine Bewegungsbeurteilung erfolgreich ist, dann lassen sich daraus Handlungsmaximen für die Trainierbarkeit der Beurteilungsprozesse sportlicher Bewegungen ableiten. Diese könnten verwendet werden, wenn es darum geht, Ausbildungsinhalte und -methoden in der Trainer- oder Sportlehrerausbildung festzulegen. Das Beurteilen sportlicher Bewegungen stellt jedenfalls sportartübergreifend eine wichtige Handlungskompetenz von Sportlehrkräften oder Trainern dar, welche im Unterrichts- und Trainingsalltag täglich vorkommt: Bewegungen beobachten und beurteilen stellt somit eine diagnostische Kernkompetenz der genannten Berufsgruppe dar.

Nach der theoretischen Beschäftigung mit der Bewegungsbeobachtung und Bewegungsbeurteilung (siehe Kapitel 2) lässt sich festhalten, dass das Wissen über Bewegungsbeurteilungsprozesse relativ gering ist. Es ist nicht viel bekannt darüber, was sinnesphysiologisch bei der Bewegungsbeobachtung abläuft oder warum ein Trainer oder Sportlehrer das Bewegungssehen gut oder schlecht beherrscht. Im Rahmen dieser Arbeit wurde daher versucht, die Fähigkeit der Beurteilung sportlicher Bewegungen als Teil der Diagnosekompetenz einer Sportlehrkraft zu definieren, welche im bereits erwähnten Begriff der fachdidaktischen Diagnosekompetenz zusammengefasst ist. Die Kompetenz zum Beurteilen sportlicher Bewegungen beruht dabei auf der Fähigkeit, schnell ablaufende Bewegungen wie die leichtathletische Schlagwurfbewegung anhand vorgegebener Bewegungsmerkmale beobachten zu können (Beobachtungskompetenz). Um diese sehr komplexe Beurteilungskompetenz messen zu können, müssen Situationen konstruiert werden, in denen die Fähigkeit zur Bewegungsbeurteilung zur Anwendung kommt. Da die Kompetenz an sich nicht messbar ist, kann nur über das in einer Beurteilungsaufgabe erzielte Ergebnis auf die Kompetenzausprägungen rückgeschlossen werden. Wenn man von konkreten Situationen im Unterrichtsalltag einer Sportlehrkraft im Zusammenhang mit dem Beurteilen sportlicher Bewegungen ausgeht, so ist das Beurteilen einer Bewegungsdemonstration sowohl zur Notengebung als auch im Sinne der Technikverbesserung eine zentrale Berufsanforderung. Dabei ist festzuhalten, dass die Beurteilungssituationen sehr vielfältig und unterschiedlich schwierig sein können. Das Rückmelden eines einzelnen Bewegungsmerkmals ist nicht so anspruchsvoll wie das Identifizieren einer abweichenden Ausführung, wenn das zu beobachtende Bewegungsmerkmal vorher nicht bekannt ist. Daher ist das Kernstück dieser Studie die Konzeption eines Beurteilungstests, der die Beurteilungskompetenz in ihren immer anspruchsvoller werdenden Facetten möglichst operationalisierbar abbildet (siehe Kapitel 4.5). Die von Meinel in die Bewegungswissenschaft eingebrachten Bewegungsmerkmale Bewegungsharmonie, Bewegungstempo, Bewegungsrhythmus, Bewegungsumfang, Bewegungskopplung und Bewegungspräzision dienen dabei als Beurteilungskriterien, um die Qualität einer Bewegungsausführung zu beschreiben.

Wenn man davon ausgeht, dass das entwickelte Messinstrumentarium valide und reliabel die Beurteilungskompetenz misst, so besteht der nächste Schritt darin, mögliche Einflussfaktoren auf die Beurteilungsleistung zu überprüfen. Wie im Theorieteil nachvollziehbar geschildert wurde, werden dafür das Fachwissen über die zu beurteilende Bewegung, die allgemeine Fähigkeit zum Ist-Sollwert-Vergleich, die Qualität der Eigenrealisation und die Vorerfahrung im Beurteilen von Bewegungen als mögliche Einflussfaktoren begründet ausgewählt. Andere Größen wie die Beobachtungsstrategie oder das dynamisch-visuelle Sehvermögen einer beurteilenden Person müssen zwar ebenfalls angenommen werden, bleiben aber in dieser Studie unberücksichtigt, weil bislang kein eindeutiger Nachweis erbracht werden konnte, dass ein eindeutiger Einfluss auf die Beurteilungsleistung gegeben ist. Somit werden die vier Variablen Fachwissen, Eigenrealisation, Fähigkeit zum Ist-Sollwert-Vergleich sowie Beurteilungsvorerfahrung auf ihre Beziehung zur Beurteilungsleistung überprüft (siehe Abb. 30).

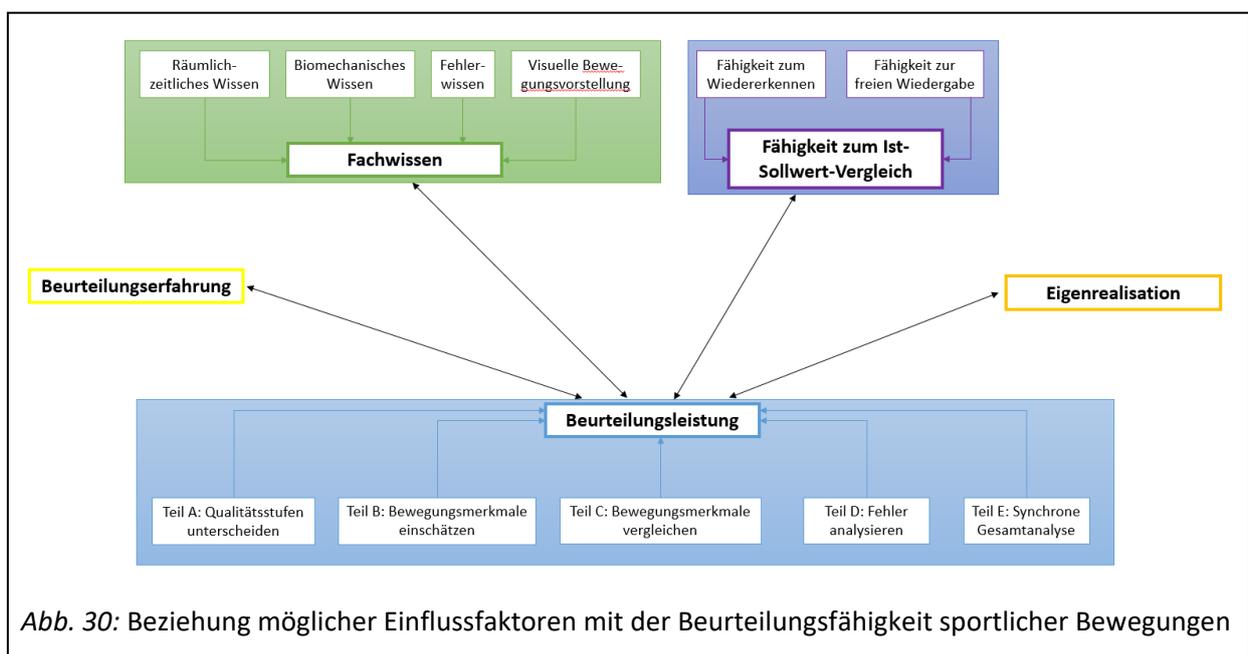


Abb. 30: Beziehung möglicher Einflussfaktoren mit der Beurteilungsfähigkeit sportlicher Bewegungen

Aus diesen vier Variablen sowie der Beschaffenheit der unterschiedlichen Beurteilungssituationen sollen die Hypothesen für den folgenden empirischen Teil dieser Studie gebildet werden. Aufgrund der wenigen Forschungsergebnisse im Bereich der Bewegungsbeurteilung bleibt aber festzuhalten, dass die gebildeten Hypothesen als wenig theoretisch abgesichert gelten müssen. Dementsprechend vorsichtig und relativ allgemein müssen die Hypothesen formuliert werden, man kann auch von Forschungshypothesen sprechen. Es empfehlen sich daher auch eher einseitige Hypothesen, weil die bisherigen empirischen Befunde nicht ausreichen, eine bestimmte Richtung in der Beziehung der beiden Variablen als relativ wahrscheinlich anzunehmen (Bös, Hänsel & Schott, 2000, S. 27). Es ist zu erwarten, dass im Rahmen dieser Studie Erkenntnisse gewonnen werden, die zur Bildung neuer Hypothesen beitragen können.

Die Beziehung der Beurteilungsleistung zum Fachwissen muss differenziert betrachtet werden. Es ist anzunehmen, dass ein hohes Fachwissen positiv mit dem Beurteilen sportlicher Bewegungen korreliert. Aufgrund der vorher aufgezeigten Erkenntnisse liegt sogar die Vermutung nahe, dass ein vertieftes biomechanisches Wissen über die Bewegung eine stärkere Vorhersagekraft für die Beurteilungs-

leistung hat als die Kenntnis des räumlich-zeitlichen Bewegungsablaufs. Das Wissen über Bewegungsabweichungen (Fehlerwissen) sollte einen ebenso positiven Zusammenhang mit dem Beurteilungsprozess aufweisen wie die Kenntnis der richtigen Reihenfolge der Schlüsselpositionen (visuelle Bewegungsvorstellung). Je anspruchsvoller die Beurteilungsaufgabe gestaltet ist, desto positiver wird der Zusammenhang mit dem Fachwissen ausfallen. Daraus lassen sich folgend Hypothesen ableiten, wobei vorausgesetzt wird, dass sich die allgemeinen Formulierungen ausschließlich auf die leichtathletische Schlagwurfbewegung beziehen:

Grundhypothese 1: Das Fachwissen korreliert positiv mit der Beurteilungsfähigkeit sportlicher Bewegungen.

- *Erweiterungshypothese 1.1: Das biomechanische Wissen steht in einem stärkeren positiven Zusammenhang mit der Beurteilungsleistung als das räumlich-zeitliche Wissen.*
- *Erweiterungshypothese 1.2: Je komplexer die Beurteilungsaufgabe ist, desto ausgeprägter ist der Zusammenhang mit dem Fachwissen.*
- *Erweiterungshypothese 1.3: Ein umfangreiches Fehlerwissen begünstigt die richtige Beurteilung der Bewegungsqualität.*
- *Erweiterungshypothese 1.4: Die visuelle Bewegungsvorstellung der Schlüsselpositionen korreliert positiv mit der Beurteilungsleistung.*

Im Bezug auf die allgemeine Fähigkeit zum Ist-Sollwert-Vergleich, welche mittels des Tests zur Kontrolle der Bewegungsvorstellungsfähigkeit (TKBV) erhoben wird, ist zu vermuten, dass diese bei einfacheren Beurteilungsaufgaben ein Indikator für die Beurteilungsleistung sein kann. Die Fähigkeit, sich einfache Bewegungsabfolgen von Körperteilen vorstellen und merken zu können, stellt eine Grundvoraussetzung für das Beobachten und Beurteilen speziellerer Bewegungen dar. Je ausgeprägter die genannten Fähigkeiten zum Ist-Sollwert-Vergleich sind, desto besser gelingen also einfache Bewegungsbeurteilungen wie das Einschätzen eines einzelnen Bewegungsmerkmals (Teil B des Beurteilungstests). Wenn die Beurteilungsaufgabe komplexer wird, wie dies beispielsweise bei Teil D des Beurteilungstests im Rahmen der Fehleridentifikation der Fall ist, dann werden Einflussfaktoren wichtiger, die auf die spezielle Bewegungsausführung gerichtet sind, und die Bedeutung der allgemeinen Bewegungsvorstellungsfähigkeit wird geringer. Da die Aufgaben des TKBV als relativ einfach eingeschätzt werden, gilt es außerdem zu überprüfen, ob sich durch schwierigere Aufgaben zum Ist-Sollwert-Vergleich die Zusammenhänge mit den Aufgaben des Beurteilungstests ändern. Hierfür wurden die Testaufgaben des TKBV dahingehend modifiziert, dass zwei gefilmte Bewegungsabfolgen präsentiert und miteinander verglichen werden mussten (siehe Kapitel 4.3.3). Es wird vermutet, dass die Ergebnisse dieses Bewegungsbeobachtungstests vor allem mit schwierigeren Beurteilungsaufgaben wie dem Bewegungsvergleich in Teil D des Beurteilungstests einen stärkeren positiven Zusammenhang zeigen als der TKBV.

Daraus können sich folgende Hypothesen ableiten:

Grundhypothese 2: Die allgemeine Fähigkeit zum Ist-Sollwert-Vergleich korreliert positiv mit der Beurteilungsfähigkeit sportlicher Bewegungen.

- *Erweiterungshypothese 2.1: Je komplexer die Beurteilungsaufgabe wird, desto geringer wird der Einfluss der allgemeinen Bewegungsvorstellungsfähigkeit.*
- *Erweiterungshypothese 2.2: Die Ergebnisse des Bewegungsbeobachtungstests korrelieren positiver mit den Ergebnissen von Teil D des Beurteilungstests als die Ergebnisse des TKBV.*

Die Eigenrealisation wird als durchgängig positiver Einflussfaktor auf die Beurteilungsleistung angenommen. Wer eine Bewegung selbst gut demonstrieren kann, hat durch die inneren Bewegungseindrücke auch einen Vorteil bei der Beurteilung von Fremdbewegungen. Möglicherweise kann auch ein gewisses Qualitätsniveau der eigenen Bewegungsausführung identifiziert werden, oberhalb dem der Einfluss der Qualität der Eigenrealisation auf die Beurteilungsleistung nachlässt. Außerdem wird begründet vermutet, dass die Eigenrealisation für komplexere Beurteilungsaufgaben eine wichtigere Rolle spielt als für einfache Beurteilungsaufgaben. Wenn aufgrund der komplexeren Beurteilungssituation die Zeit sowie die Aufmerksamkeit zum Beobachten einzelner Bewegungsmerkmale geringer wird, so ist ein Beurteiler mit einer besseren Eigenrealisation im Vorteil, weil er die Fremdbewegung besser nachempfinden und mit seinen eigenen Bewegungserfahrungen vergleichen kann, auch wenn kein bewusstseinsfähiger Bewegungsvergleich möglich ist.

Grundhypothese 3: Je besser die Qualität der Eigenrealisation ist, desto besser kann eine Bewegungsbeurteilung durchgeführt werden.

- *Erweiterungshypothese 3.1: Mit Zunahme der Qualität der Eigenrealisation nimmt der Einfluss auf die Beurteilungsleistung ab.*

Die Vorerfahrung im Beurteilen sportlicher Bewegungen soll sich auch positiv auf die Beurteilungsleistung auswirken. Je ähnlicher die bereits beurteilten Bewegungen zur zu beurteilenden Zielbewegung sind, desto positiver zeigt sich der Transfereffekt. Ebenso ist anzunehmen, dass bei zunehmend komplexer werdenden Beurteilungen die Spezifität der Beurteilungsvorerfahrung eine stärkere Rolle spielt.

Grundhypothese 4: Je spezifischer die Beurteilungsvorerfahrung hinsichtlich der zu beurteilenden Bewegung ausfällt, desto positiver ist der Zusammenhang mit der Beurteilungsleistung.

4 Methodik

Die methodische Schwierigkeit dieser Studie besteht darin, dass es noch kein empirisch überprüftes Modell für die Erfassung der Beurteilungsfähigkeit sportlicher Bewegungen gibt. Somit sind auch die Einflussfaktoren auf die Beurteilungsleistung noch weitgehend ungeklärt. Dies ist keineswegs überraschend, denn in der Kompetenzforschung besteht über wesentlich grundlegendere Konstrukte wie das Professionswissen und dessen Modellierung noch weitgehend Uneinigkeit. Um mögliche Einflussfaktoren auf die Beurteilungsleistung quantitativ zu untersuchen, müssen daher möglichst valide und reliable Testinstrumente entwickelt werden. Angesichts der fehlenden Forschungsbemühungen ist nicht unbedingt damit zu rechnen, dass signifikante Abhängigkeiten der Beurteilungsleistung von den genannten Einflussfaktoren festgestellt werden können, aber möglicherweise werden gewisse Tendenzen identifiziert, die weiter erforscht werden können. Es ist durchaus möglich, dass ein detailliertes Fachwissen über die Wurfbewegung keine Voraussetzung für die Beurteilungsleistung darstellt, obwohl sich dies vermuten lässt.

Aus diesem Grund müssen zum Zwecke dieser Studie die meisten Erhebungsinstrumente selbst konstruiert werden. Das Fachwissen über die leichtathletische Wurfbewegung (siehe Kapitel 4.2) wird ebenso durch einen eigenen Multiple-Choice-Test erhoben wie die Fähigkeit, Wurfbewegungen von Schülerinnen und Schülern qualitativ zu beurteilen (siehe Kapitel 4.5). Der bereits bestehende TKBV-Test von Schott (2013) kommt bei der Überprüfung der Fähigkeit zum Ist-Sollwert-Vergleich zum Einsatz und wird durch einen eigenen Testteil (siehe Kapitel 4.3) erweitert, bei dem vor allem die Fähigkeit der Bewegungsbeobachtung und des Bewegungsvergleichs im Mittelpunkt stehen. Zur Einschätzung der Eigenrealisation wird anhand von Videoaufzeichnungen, die qualitativ und quantitativ analysiert werden, ein Index über die Qualität der Bewegungsdemonstration erstellt (siehe Kapitel 4.4).

Im Folgenden wird dargestellt, welche Stichprobe der Studie zugrunde liegt und wie die Erhebungsinstrumente inhaltlich entwickelt werden, um den wissenschaftlichen Gütekriterien zu entsprechen. Als Anhaltspunkt für die methodische Entwicklung der einzelnen Testverfahren dient das Konstruktionschema von Bös (siehe Abb. 31), welches für die Konstruktion motorischer Tests schon seit mehreren Jahrzehnten zur Verfügung steht (Bös & Wottawa, 1987). Es geht dabei nicht darum, alle aufgeführten Schritte monoton zu bearbeiten, sondern bei der Konstruktion neuer Testverfahren geordnet vorzugehen und die einzelnen Schritte sinnvoll zu reflektieren. Die erste Frage nach der Relevanz der Anwendung von Testverfahren für die Messung der Beurteilungsleistung und ihrer möglichen Einflussfaktoren wurde bereits in Kapitel 3 ausführlich beschrieben. Auch die fehlende Existenz geeigneter Testverfahren für die zu untersuchenden Variablen wurde abgesehen vom Test der Kontrolle der Bewegungsvorstellungsfähigkeit (TKBV) zur Messung der allgemeinen Fähigkeit zum Ist-Sollwert-Vergleich bereits als spezielle Problemstellung der Beschäftigung mit der Beurteilungskompetenz dargelegt.

Vor der eigentlichen Testkonstruktion sollen dabei zuerst die Testziele formuliert werden, ehe Gedanken über die inhaltlichen Grundlagen, die Konstruktionsstrategie und die testtheoretischen Grundlagen des Testkonzepts angestellt werden. Die Festlegung des Aufgabentyps und der Bewertungskriterien sowie die Operationalisierung der Testaufgaben führen die theoretischen Überlegungen in praxisbezogene Entscheidungen und Formulierungen der Testitems über. Als **testtheoretisches Modell** wird für alle in dieser Studie verwendeten Testverfahren die probabilistische Item-Response-Theorie (siehe Kapitel 4.7.1) gewählt, wenngleich Bös (2017b, S. 817) darauf verweist, „dass die messmethodisch

schärferen probabilistischen Testmodelle vor allem erfolgversprechend sind, wenn der Merkmalsbereich bereits genügend erforscht und strukturiert ist“. Die empirisch-statistischen Arbeitsschritte der Aufgabenanalyse, Aufgabenselektion sowie Testanalyse stehen am Ende der Testkonstruktion und werden im Ergebnisteil dieser Studie (siehe Kapitel 5) entsprechend ausgewertet. Die letzten Phasen von der Testendform bis zur Testevaluation werden im Rahmen dieser Untersuchung nicht mehr angewendet. Mithilfe dieses Konstruktionsalgorithmus soll eine hohe methodische Güte der entsprechenden Testverfahren sichergestellt werden (siehe Abb. 31), um vor allem dem Repräsentationsproblem zu begegnen, dass die bestehende Wirklichkeit durch das Messinstrument adäquat abgebildet wird (Bös & Wottawa, 1987, S. 109f.).

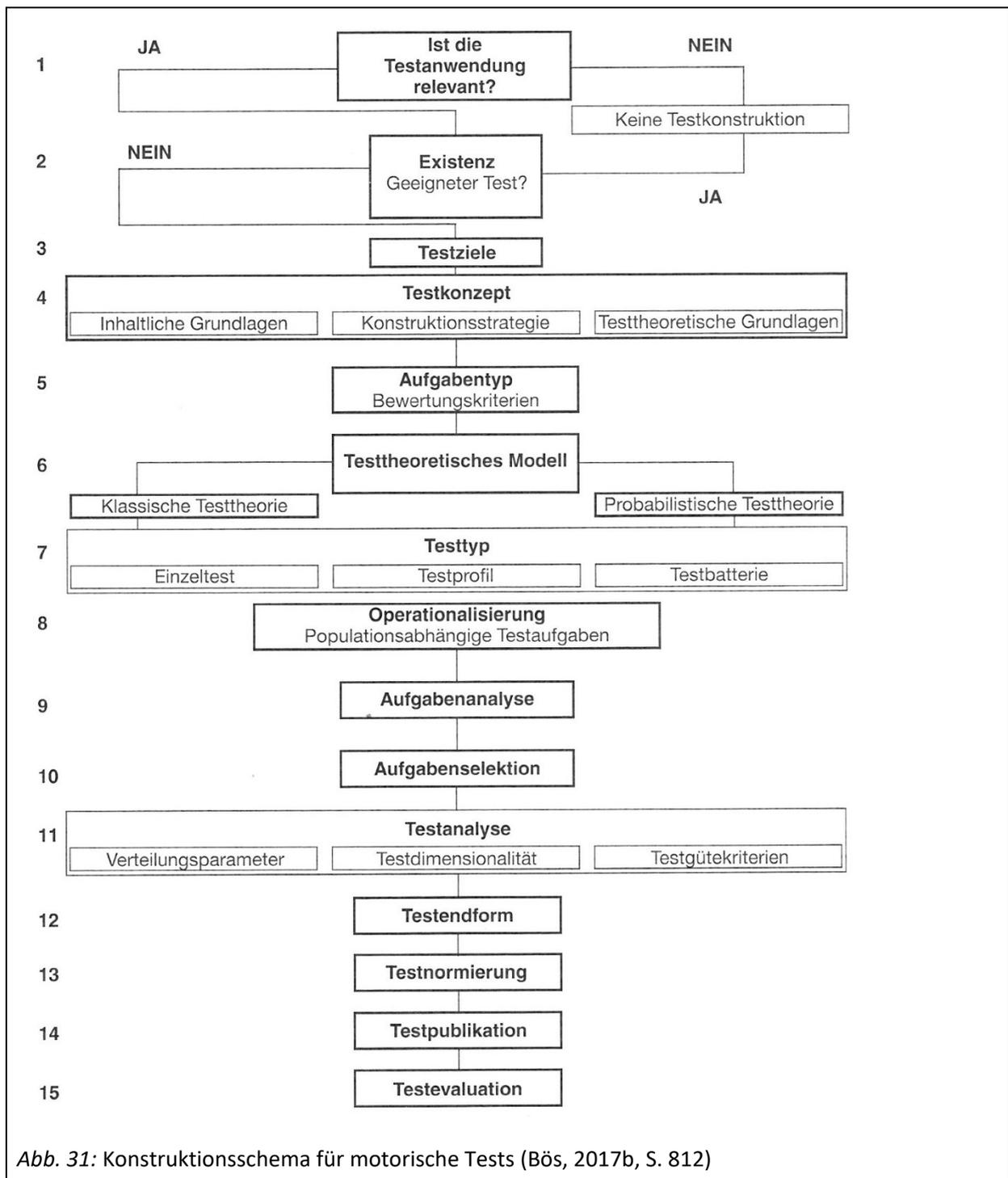


Abb. 31: Konstruktionschema für motorische Tests (Bös, 2017b, S. 812)

4.1 Testpersonen

Die Testpersonen dieser Studie rekrutierten sich aus Lehramtsstudierenden mit Hauptfach Sport an der Universität Passau. Insgesamt nahmen 59 Studierende (35 männlich, 24 weiblich) daran teil, wobei 42 von ihnen zugleich Teilnehmer des Seminars Bewegungswissenschaft des Sommersemesters 2017 waren. Die Auswahl der Probandinnen und Probanden ist somit als nichtprobabilistische Gelegenheitsstichprobe zu bezeichnen. Aus forschungsökonomischen Gründen ist das Vorgehen vertretbar, die Testverfahren im Rahmen eines sportwissenschaftlichen Seminars durchzuführen. Zum einen fügten sich die Testverfahren inhaltlich in das Seminarthema Bewegungsanalyse ein, zum anderen kann auf diese Weise eine relativ große Zahl an Personen getestet werden, die ansonsten nicht die Zeit gefunden hätten, an der aufwändigen Studie teilzunehmen. Die Testverfahren verlangen nämlich von den Testpersonen insgesamt eine zeitliche Präsenz von 240–300 Minuten. Zusätzlich folgten noch 17 Lehramtsstudierende mit Hauptfach Sport der per Email an alle vertieften und nicht-vertieften Sportstudierenden verschickten Einladung, an der Studie teilzunehmen. Die Datenerhebungen erfolgten zum Teil getrennt, zum Teil gemeinsam oder gemischt mit den Teilnehmerinnen und Teilnehmern des Seminars Bewegungswissenschaft.

4.2 Fachwissen über die Ausführung von Wurfbewegungen

Das Ziel eines Testverfahrens, welches das fachliche Wissen über Bewegungsausführungen misst, besteht darin, den individuellen Ausprägungsgrad des Fachwissens bei allen Testpersonen möglichst objektiv quantitativ zu erfassen. Das Kompetenzmodell von Baumert und Kunter (2006) geht davon aus, dass sich professionelle Handlungskompetenz von Lehrkräften durch Wissens- und Könnensleistungen (oder anders ausgedrückt deklaratives, prozedurales und strategisches Wissen) auszeichnet. Orientiert an Shulman (1987, S. 8) bilden das allgemeine pädagogische Wissen, das Fachwissen und das fachdidaktische Wissen die zentralen Wissens- und Könnensbereiche des sog. Professionswissens (siehe Abb. 32).

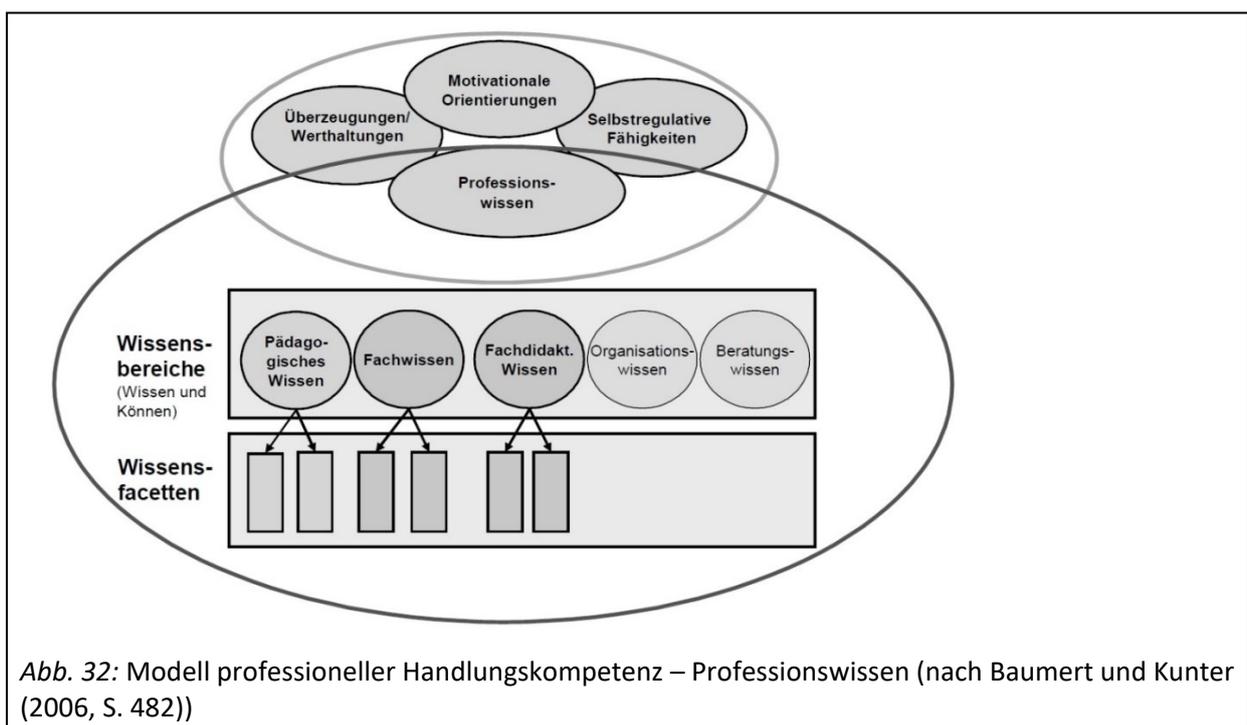


Abb. 32: Modell professioneller Handlungskompetenz – Professionswissen (nach Baumert und Kunter (2006, S. 482))

Obwohl Fachwissen und fachdidaktisches Wissen unbestritten zum Kern professioneller Lehrkompetenz gehören und in der universitären Lehramtsausbildung im Mittelpunkt stehen, so wird beides in der Diskussion über das praktische Handlungswissen von Lehrpersonen geradezu geringgeschätzt oder übersehen. Dabei kritisiert Terhart (2002, S. 24), dass fachbezogenes Wissen und Kompetenz nicht zu den Standards zählen, denn „das Verfügen über ein gutes fachbezogenes Wissen [sei] doch sicherlich eine notwendige und sehr wichtige (allerdings noch nicht hinreichende!) Voraussetzung für kompetenten Unterricht in diesem Fach“. Oser und Renold (2005, S. 126) hingegen betonen, dass der Wissensbegriff bezüglich seiner Vorhersagekraft für erfolgreiches Lehrerhandeln begrenzt ist:

Alle wissen, dass man mit dem Abrufen und Testen der ‚knowledge bases‘ von Lehrpersonen fast nichts über das Verhalten in der Situation aussagen kann, und trotzdem messen wir nichts anderes als diese Wissensbereiche, die meistens fachbezogen oder allgemein erziehungswissenschaftlich resp. didaktisch sind.

Auch wenn es sich zunächst um vermeintlich träges Wissen handeln mag, so ist eine detaillierte, leibbildhafte Bewegungsvorstellung einer sportlichen Bewegung eine wichtige Voraussetzung für einen Beobachter, um den nur sehr zeitbegrenzt möglichen Soll-Istwert-Vergleich in einer realen Beurteilungssituation durchzuführen. Vor allem bei Bewegungen, die mit sehr hoher Geschwindigkeit ablaufen (wie die leichtathletischen Sprung- oder Wurfbewegungen) und aufgrund ihres azyklischen Verlaufs nur einmal betrachtet werden können, ist das Erfassen der Bewegungsqualität daher besonders anspruchsvoll.

Wie in Abb. 32 angedeutet, lässt sich das Fachwissen in weitere Wissensfacetten unterteilen, die je nach Wissensinhalt individuell inhaltlich geklärt werden müssen. Dabei kann zum einen das Wissen über den räumlich-zeitlichen Verlauf eine erste Kenngröße für die Qualität der Bewegungsvorstellung sein. Zum anderen können auch vertiefte Kenntnisse aus der Biomechanik oder Wissen über häufig auftretende Mängel bei der Bewegungsausführung hilfreich sein, um die Bewegungsbeobachtung zu unterstützen. Ebenso gilt es zu überprüfen, wie gut das visuelle Vorstellungsvermögen der Wurfbewegung anhand von Schlüsselpositionen ausgeprägt ist. Für die Messung dieser Wissensbereiche gilt es entsprechend aussagekräftige Testverfahren zu entwickeln, das den wissenschaftlichen Gütekriterien entspricht.

Aus den Erkenntnissen über die Strukturierungen des Wissensbereichs in der Fachliteratur wird ein **Testkonzept** entwickelt, welches die einzelnen Facetten des Wissensbereichs darstellt (siehe Abb. 33). Auf den Grundlagen dieser konzeptionellen Überlegungen werden vier Elementartests konstruiert, welche die einzelnen Wissensfacetten repräsentieren. Die vier Einzeltests behalten im Rahmen dieser Studie ihre Eigenständigkeit, werden aber bei der Testdurchführung zu einem **Testprofil** zusammengefasst, um sicherzustellen, dass die ausgewählten Anforderungsbereiche zielgerichtet und gleichmäßig mit Aufgaben abgedeckt werden (siehe Schritt 7 des Konstruktionsalgorithmus in Abb. 31).

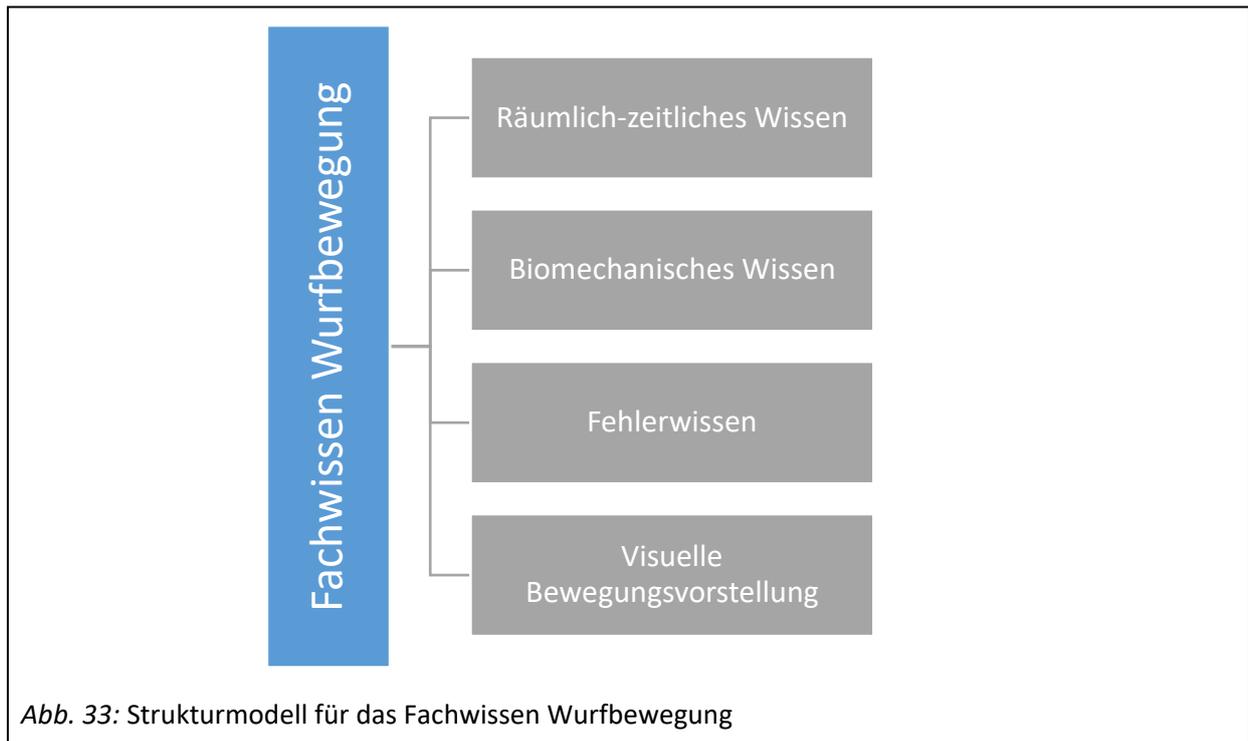


Abb. 33: Strukturmodell für das Fachwissen Wurfbewegung

Der Fachwissenstest wird demnach in folgende vier Elementartests untergliedert:

- Teil A: Wissen über räumlich-zeitlichen Bewegungsverlauf (RZ-Wissen)
- Teil B: Biomechanisches Wissen
- Teil C: Fehlerwissen
- Teil D: Visuelle Bewegungsvorstellung

Als nächstes folgen Überlegungen und Entscheidungen über die **Wahl des Aufgabentyps** zur Überprüfung der einzelnen Wissensfacetten entsprechend vorgegebener Bewertungskriterien (siehe Schritt 5 des Konstruktionsalgorithmus in Abb. 31). Die Wissenserfassung erfolgt in Teil A und Teil B in einem gebundenen Antwort-Wahl-Verfahren, in welchem die Probandinnen und Probanden für jedes Item entscheiden, ob dieses richtig oder falsch ist. Es handelt sich somit um dichotome Daten als Spezialfall einer Nominalskala mit nur zwei möglichen Ausprägungen, was bei der statistischen Datenauswertung später berücksichtigt werden muss. In Teil C kommt das freie Antwortformat zur Anwendung, in dem eine Testperson auf die gestellte Aufgabe, ihr Wissen über häufig auftretende Abweichungen bei der Wurfbewegung aufzuschreiben, mit einer mehr oder weniger detaillierten selbstgewählten Ausführlichkeit antwortet (Kubinger, 2011, S. 5f.). Teil D wird als Sortieraufgabe durchgeführt, in welcher verschiedene Bilder der Wurfbewegung in die richtige Reihenfolge gebracht werden müssen. Bei den Teilen C und D handelt es sich somit um eine Kardinalskala mit diskreter Merkmalsausprägung, da der Punktwert einer eindeutig quantifizierbaren, jedoch nicht weiter unterteilbaren Größe entspricht. Je nach Antwortqualität (Anzahl richtiger oder falscher Nennungen) werden den Probanden entsprechende Punktwerte zugeordnet, welche in eine natürliche Reihenfolge gebracht werden können. Auch die Abstände zwischen den einzelnen Bewertungseinheiten sind eindeutig interpretierbar, da jede Bewertungseinheit dem Vorhandensein einer richtigen Antwort entspricht.

Bei der Testdurchführung müssen die Probandinnen und Probanden zuerst Teil C (Fehlerwissen) absolvieren, bevor Ihnen der Testbogen für Teil A und Teil B (diese beiden Teile waren auf einem Testbogen gemeinsam abgebildet) ausgehändigt wird. Dies sollte gewährleisten, dass das Fehlerwissen wirklich aus dem Gedächtnis abgerufen und nicht durch die vorherige Auseinandersetzung mit den gebundenen Items beeinflusst wird. Nach Abgabe des Testbogens von Teil A und B wird den Testpersonen Teil D (Visuelle Bewegungsvorstellung) ausgehändigt. Der Test wurde ohne Zeitbeschränkung durchgeführt, die Bearbeitungsdauer lag bei allen Probandinnen und Probanden zwischen 45 und 60 Minuten.

Die **Operationalisierung** der Wissensfacetten (siehe Schritt 8 des Konstruktionsalgorithmus in Abb. 31) erfolgt mittels Top-Down-Technik, indem aus der bestehenden Fachliteratur mögliche Inhaltsbereiche identifiziert sowie Indikatoren und Definitionsmerkmale eruiert werden. Die Auswahl möglicher Fachliteratur gestaltet sich in der Leichtathletik überschaubar, da es nur wenige aktuelle Standardwerke gibt, welche die Wurfbewegung in einer Detailliertheit beschreiben, wie sie für die Entwicklung eines wissenschaftlich anspruchsvollen Fachwissenstests nötig sind. Folgende Publikationen kommen dabei zum Einsatz:

- Bauersfeld, K.-H. & Schröter, G. (2016). *Grundlagen der Leichtathletik: Das Standardwerk für Ausbildung und Praxis* (6., aktualisierte Auflage). Aachen: Meyer & Meyer Verlag.
- Strüder, H. K., Jonath, U. & Scholz, K. (2016). *Leichtathletik: Trainings- und Bewegungswissenschaft - Theorie und Praxis aller Disziplinen* (2. korrigierte Auflage). Hellenthal: Sportverlag Strauß.
- Killing, W. (Hrsg.). (2011). *Jugendleichtathletik Wurf: Offizieller Rahmentrainingsplan des Deutschen Leichtathletik-Verbandes für die Wurfdisziplinen im Aufbautraining; [Wurf]*. Münster: Philippka-Sportverlag.
- Ballreich, R. & Kuhlow, A. (Hrsg.). (1986). *Biomechanik der Leichtathletik. Biomechanik der Sportarten: Vol. 1*. Stuttgart: Enke.

In allen Standardwerken nimmt dabei die Beschreibung der Wurftechnik nach räumlich-zeitlichen Kriterien eine zentrale Position ein. Überblicksgraphiken über die Phasenstruktur des Speerwurfs wie bei Bauersfeld und Schröter (2016, S. 594f.) oder Strüder, Jonath und Scholz (2016, S. 685f.) geben einen ersten Eindruck, dass viele Detailinformationen erforderlich sind, um die Speerwurfbewegung umfassend zu verstehen. Eine Zusammenfassung der wichtigsten räumlich-zeitlichen Bewegungskriterien ist in Kapitel 4.2.1.1 zu finden. Aus diesen Einzelinformationen werden Items für Teil A (RZ-Wissen) konstruiert, mithilfe derer das Fachwissen über den räumlich-zeitlichen Verlauf der Wurfbewegung erfasst werden soll.

Das Biomechanische Wissen wird in der Überblicksliteratur am ausführlichsten von Strüder et al. (2016) behandelt. Zahlreiche Tabellen mit biomechanischen Messwerten oder Diagramme über Kraft-Zeit-Verläufe deuten auf die Wichtigkeit biomechanischer Kenntnisse für das Gesamtverständnis der Wurfbewegung hin. Bauersfeld und Schröter (2016) nutzen biomechanische Untersuchungsergebnisse hauptsächlich dafür, das Verständnis für räumlich-zeitliche Bewegungsmerkmale zu verbessern, indem immer wieder Winkel- und Längenangaben oder Geschwindigkeitswerte zur Veranschaulichung herangezogen werden. Ballreich und Kuhlow (1986) beschäftigt sich mit der Biomechanik der Leichtathletik sogar in einem eigenen Buch und behandelt die biomechanischen Einflussgrößen der einzelnen Phasen eines Speerwurfs. Außerdem gibt es mittlerweile zahlreiche Publikationen über biomechanische Messungen bei vergangenen Leichtathletik-Weltmeisterschaften, in welchen die Würfe der Finalistinnen und Finalisten auf unterschiedliche Parameter hin analysiert werden (z. B. Murakami et al. (2006) oder Lehmann (2010)). Es wird demnach deutlich, dass biomechanisches Wissen zum einen für

die Techniko Optimierung eine wichtige Rolle spielt, zum anderen aber auch Trainern und Lehrkräften dienen kann, ein besseres Gesamtverständnis der leichtathletischen Wurfbewegung zu erlangen. Eine Zusammenfassung der biomechanischen Kenntnisse anhand einschlägiger Fachliteratur wird in Kapitel 4.2.2.1 gegeben. Aus diesen Erkenntnissen werden dann biomechanische Aussagen generiert, welche die Probandinnen und Probanden im Rahmen des Fachwissenstests als richtig oder falsch einschätzen müssen (siehe Kapitel 4.2.2.2).

Auffallend bei der Durchsicht der Standardliteratur ist auch, dass Kapitel zu finden sind, die „Fehler-Ursachen-Korrekturen“ (Strüder et al., 2016, S. 702–706) oder „Wesentliche Fehler und Korrekturmöglichkeiten“ (Bauersfeld & Schröter, 2016, S. 615f.) heißen. Das Vermitteln von Fehlerbildern oder Abweichungen vom idealen Bewegungsverlauf sowie das Wissen um Ursachen und Korrektur scheinen einen festen Platz in der fachdidaktischen Literatur zu haben. Marschall, Blischke, Müller und Reiser (1996, S. 196) sprechen von einer fehlerbezogenen Wissensbasis, welche auf den Prozess der Bewegungsbeobachtung und Fehleridentifikation Einfluss haben könnte. In einer Synopse der Fachliteratur werden wesentliche Fehlerbilder identifiziert und zusammengefasst, welche dann als Erwartungshorizont für die Auswertung der Freitextaufgabe von Teil C (Fehlerwissen) dienen (siehe Tab. 45 im Appendix).

Die Überprüfung der visuellen Vorstellungsfähigkeit der Wurfbewegung wird mittels einer Aufgabenstellung erreicht, in welcher die Probandinnen und Probanden acht unsortiert angeordnete Körperpositionen in eine dem chronologischen Bewegungsablauf der Wurfbewegung entsprechende Reihenfolge bringen müssen (siehe Kapitel 4.2.4). Die Auswertung wird entsprechend standardisiert und das Ergebnis kann als isolierte Fähigkeit zur visuellen Bewegungsvorstellung der Wurfbewegung verwendet werden oder in den Gesamtscore des Fachwissens mit einfließen. Mit dem Gesamtwert sowie den Einzelergebnissen der jeweiligen vier Teile wird dann versucht, die Bedeutung des Fachwissens auf die Beurteilungsleistung von Wurfbewegungen in Beziehung mit anderen Einflussfaktoren zu setzen.

4.2.1 Wissen über den räumlich-zeitlichen Verlauf der Wurfbewegung

Als **inhaltliche Grundlage der Testkonzeption** (siehe Schritt 4 des Konstruktionsalgorithmus in Abb. 31) von Teil A (RZ-Wissen) wird mittels genannter leichtathletischer Standardliteratur eine detaillierte Sollwertbestimmung der leichtathletischen Wurfbewegung erstellt, welche sich überwiegend an Aussagen zum Speerwurf orientiert, da dieser „die wesentlichste Wettkampfdisziplin aus der Gruppe der Schlagwürfe [darstellt], zu der neben dem Speerwurf auch der Schlagballwurf gehört“ (Bauersfeld & Schröter, 2016, S. 593). Die technischen Anforderungen im Schlagballwurf ordnen sich bewusst in das Anforderungsprofil des Speerwurfs ein, da der Ballwurf zur Ausbildung von Teilelementen der Speerwurftechnik unter erleichterten Bedingungen genutzt werden soll (Bauersfeld und Schröter (2016, S. 602)). Auch Killing (2011, S. 136) betont, dass Schlagballwürfe aufgrund der engen Bewegungsverwandtschaft eine besondere Bedeutung für den Speerwurf haben. Der Ballwurf selbst wird in der Literatur jedoch meist nur im Zusammenhang mit der Wurfschulung von Anfängern thematisiert, wobei in diesem Kontext auf allzu detaillierte technische Kriterien verzichtet wird, weshalb das technische Leitbild des Speerwurfs maßgebend ist für die Beurteilung von leichtathletischen Schlagwurf-bewegungen. Darum wird in Kapitel 4.2.1.1 auch der Technikanalysebogen von Tidow für den Speerwurf an geeigneten Stellen zitiert, welcher eine sehr detaillierte Auflistung von Technikelementen in den verschiedenen Bewegungsphasen beinhaltet. Dabei gilt es bei der **Aufgabenanalyse und Aufgabenselektion** (siehe Schritte 9 und 10 des Konstruktionsalgorithmus in Abb. 31) zu bedenken, dass

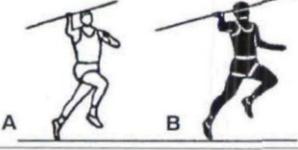
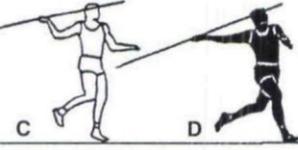
gemäß der Hierarchisierung von critical features (siehe Kapitel 2.3.2.1) nur die Technikdetails für die Bewegungsbeobachtung eine Rolle spielen, die eine wichtige Funktion für eine gelungene Bewegungsausführung erfüllen. In einem zweiten Schritt erfolgt dann eine sinnvolle Auswahl dieser Technikdetails, um nur solche Kriterien in die Beschreibung der Bewegungsmerkmale aufzunehmen, die auch einer altersgemäßen Bewegungsausführung des leichtathletische Schlagwurfs entsprechen (siehe Kapitel 2.3.2.4). Zur besseren Nachvollziehbarkeit der Itemkonstruktion werden im nächsten Kapitel bei den entsprechenden Sollwertformulierungen bereits Verweise auf den zur Erfassung des Fachwissens erstellten Fragebogen hergestellt.

4.2.1.1 Sollwertbestimmung der Wurfbewegung

Als Orientierungsgrundlage und Visualisierungshilfe der folgenden Sollwertbestimmung dient der Technikanalysebogen von Tidow (1996a). Die Phasenstruktur der Schlagwurfbewegung kann gemäß Strüder et al. (2016, S. 668) grundsätzlich in den zyklischen und azyklischen Anlaufteil, die Abwurfbewegung und das Abfangen unterschieden werden, wobei Bauersfeld und Schröter (2016, S. 594f.) die Abwurfphase mit dem ersten Bodenkontakt des Druckbeins nach dem Impulsschritt beginnen und mit dem Verlassen des Wurfgeräts von der Wurfhand enden lassen. Die Anlaufphase wiederum kann in einen zyklischen und einen azyklischen Teil unterschieden werden. Der **zyklische Anlauf** dient als lockerer Steigerungslauf (bei Fortgeschrittenen meist über 8–12 Schritte) der Beschleunigung des Gesamtsystems Werfer/Gerät, wobei noch kein betontes Vor- und Rückführen des Wurfarms erfolgt (siehe Tab. 6, Kriterien A1 und A3). Becken, Brust und Kopf sind frontal in Wurfrichtung ausgerichtet, der Rumpf ist am Ende des zyklischen Anlaufteils möglichst aufrecht (Bauersfeld & Schröter, 2016, S. 596).

Als **azyklischer Anlaufabschnitt** werden die letzten vier Bodenkontakte vor dem Aufsetzen des Druckbeins verstanden (5er-Rhythmus), durch welche der Körper in eine optimale Abwurfposition gebracht werden soll. Zu Beginn dieser Phase erfolgt die Rückführung des Wurfgeräts, indem der Wurfarm synchron mit einer Drehung der Wurf Schulter um 90° nach hinten gestreckt wird (siehe Tab. 6: Kriterien BD6 und D10). Die Rückführung kann dabei allmählich über zwei bis drei Anlaufschritte erfolgen oder nahezu vollständig auf einem Schritt (entweder als Passgangbewegung auf dem ersten Schritt des 5er-Rhythmus oder in der üblichen Kreuzkoordination auf dem zweiten Schritt des 5er-Rhythmus) ausgeführt werden. Bei der schnellen Speerrückführung wird der erste Schritt des 5er-Rhythmus zumeist als weiter Druckschritt bzw. schneller flacher Sprung ausgeführt, was einen bewussten und deutlichen Übergang vom zyklischen zum azyklischen Anlaufteil begünstigt. Die Schulterachse zeigt nach der Rückführung in Wurfrichtung und auch die Hüftachse ist stark aus der Frontalposition gedreht (Achsenparallelität) (siehe Tab. 6: Kriterium D13). Durch einen aktiven, impulsgebenden Fußaufsatz sowie einen möglichst kurzen Bodenkontakt der Füße soll auch in dieser seitlichen Laufbewegung der Körper noch weiter beschleunigt werden (Bauersfeld & Schröter, 2016, S. 597). Der im Ellbogen gebeugte Gegenarm befindet sich angehoben in Wurfrichtung auf Oberkörperhöhe und der Blick geht weiter in Wurfrichtung (Strüder et al., 2016, S. 670). Aus diesen Sachinformationen zur Anlaufphase mit Rückführung des Wurfgeräts wurden vier Items für den Fachwissens-Test entwickelt, welche in Tab. 7 einzusehen sind.

Tab. 6: Technikanalysebogen für die Phasen Anlauf und Speerrücknahme (Tidow, 1996a, S. 60)

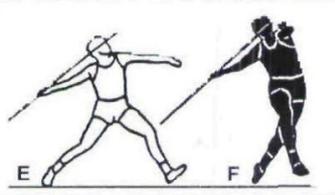
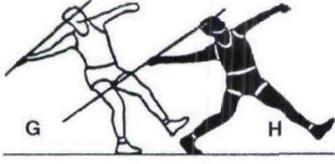
Position	Phase	Reference	Criterion
	(cyclic) Approach Start of withdrawal ↓ Withdrawal of the javelin ↓ Finish	A 1 Organization	Acceleration run
		A 2 Javelin carry	Shaft at head height / tip points in throwing direction
		A 3 Throwing arm / head	Bent / correct grip / near head
		B 4 Support foot	Contact with check mark
		BD 5 Shaft of the javelin	Gliding backward at height of temple
		BD 6 Upper body	Rotates against throwing direction
	CD 7 Foot plant	Running-adequate	
	CD 8 Trunk	Kept upright	
	CD 9 Javelin movement	Smooth / fluent backward glide	
	D 10 Throwing arm	Extended / parallel to the ground	
	D 11 Throwing hand	Turned inward / at the level of the lower arm	
	D 12 Tip of the javelin	At the height of the temple / close to hand	
	D 13 Shoulder axis	Points in throwing direction	

Tab. 7: Testvignette zur Phase der Ballrückführung (schwarz gefüllte Felder sind richtig formulierte Items, unausgefüllte Felder sind falsche Antworten)

<p>1.1 Kreuzen Sie die Aussagen an, welche bezüglich der Ballrückführung richtig sind. (4 Punkte)</p> <p><input type="checkbox"/> Die Ballrückführung erfolgt im letzten Anlaufschritt.</p> <p><input type="checkbox"/> Der Gegenarm wird mit der Ballrückführung am Körper angelegt und zeigt nach unten zum Boden.</p> <p><input checked="" type="checkbox"/> Während der Ballrückführung dreht der Athlet die Schulterachse von einer frontalen in eine parallele Position zur Wurfrichtung.</p> <p><input checked="" type="checkbox"/> Der Wurfarm wird auf Schulterhöhe zurückgeführt und dann im Ellbogengelenk gestreckt.</p>

Mit dem Aufsetzen des drittletzten Bodenkontakts beginnt der **Impulsschritt** – ein verlängerter, jedoch flach ausgeführter Schritt (siehe Tab. 8: Kriterium F19) –, an dessen Ende der Werfer mit dem Setzen des Druckbeins die charakteristische Körperposition der Wurfauslage vorbereitet. Die zurückgedrehte Hüft-, Rumpf- und Schulterposition sowie der gestreckte Wurfarm (Kriterien EF17 und EF18) bleiben während des Impulsschrittes ebenso erhalten wie die leichte Rumpfrücklage. Im Verlaufe des vorletzten Schrittes kommt es zu einem doppelten Beinwechsel (Kreuzschritt), indem beim Abdruck vom drittletzten Bodenkontakt das freie, beim Rechtswerfer rechte Bein das Abdruckbein zunächst überholt (Kriterium F20), dann vor dem Aufsetzen des rechten Druckbeins aber das freie linke Bein so schnell nach vorne beschleunigt wird, dass es das Druckbein bereits wieder überholt hat (Kriterium FG21). Die Landung des rechten Druckbeins erfolgt über die Fußaußenkante mit schräg nach vorne gerichteter Fußspitze (Kriterium G22), wobei die Fußaußensätze (mit 20–45°) deutlich weniger aus der Frontalposition gedreht sind als die Hüft- und Schulterachse (Killing, 2011, S. 145f.). Das Kniegelenk wird durch den Amortisationsdruck auf ca. 130–150° gebeugt (Kriterium G23) (Bauersfeld & Schröter, 2016, S. 598). Die Items, welche die räumlich-zeitliche Struktur des Impulsschrittes betreffen, sind in Tab. 9 dargestellt.

Tab. 8: Technikanalysebogen für die Phasen des Impulsschritts (Tidow, 1996a, S. 60)

	Intermediate stride	DE 14 Shoulder axis	Slightly diagonal position
	↓ Impulse torque	E 15 Throwing arm / javelin	Constant position as related to shoulder line
		EF 16 Foot plant	Active and emphasis on landing
		EF 17 Shoulder axis	Increasingly diagonal position
		EF 18 Throwing arm / javelin	Synchronous with shoulder movement
		F 19 Push-off	Accentuated / flat
	Support contact	G 22 Tip of support foot	Points (diagonally) in throwing direction
		G 23 Knee of support leg	"Soft" landing
	↓ Drive split position	G 24 Throwing arm / javelin	"Long" / constantly diagonal / close to head
		GH25 Support leg / knee	Rolling over ball of the foot / constant angle
		H 26 Throwing arm	Still extended / lifted parallel to the ground
		H 27 Javelin alignment	Longitudinal axis points into throwing direction
		H 28 Tip of javelin	At height of head / close to head
		H 29 Bracing leg	Heel leads / extended and pre-tensed
		H 30 Free arm	In line with shoulder axis / parallel to javelin

Tab. 9: Testvignette zur Phase des Impulsschrittes

<p>2.1 Kreuzen Sie die Aussagen an, welche bezüglich des Impulsschrittes richtig sind. (4 Punkte)</p> <p><input checked="" type="checkbox"/> Der Impulsschritt wird durch ein ziehendes Aufsetzen und einen kräftigen Abdruck im drittletzten Bodenkontakt eingeleitet.</p> <p><input type="checkbox"/> Der Impulsschritt wird mit einer hohen Flugkurve weit nach vorne ausgeführt.</p> <p><input checked="" type="checkbox"/> Während des Abdrucks zum Impulsschritt schwingt das Bein der Wurfarmseite am anderen Bein vorbei nach vorne.</p> <p><input type="checkbox"/> Der Oberkörper wird durch den Impulsschritt in eine leichte Vorlage gebracht.</p>

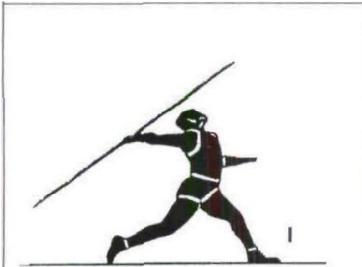
In der **Abwurfphase** werden zusätzlich zum Anlauf noch zwei Drittel der Abfluggeschwindigkeit erzeugt, weshalb in dieser kurzen Zeitspanne des Abwurfs die Hauptbeschleunigung des Wurfgeräts erfolgt. Der Körperschwerpunkt des Werfers wird vom vorletzten zum letzten Bodenkontakt möglichst ohne Geschwindigkeitsverlust in Wurfrichtung bewegt, indem eine aktive, flache Druckbewegung vom rechten Bein (beim Rechtswerfer) erfolgt. Durch das Eindrehen des rechten Fußes und ein aktives Nach-vorn-unten-Drücken des rechten Knies in Wurfrichtung (Druckbeinarbeit) wird die rechte Hüftseite nach vorne beschleunigt, während der Wurfarm noch gestreckt zurückgehalten wird (siehe Tab. 11: Kriterium HI32). Parallel dazu erfolgt das aktive, flache Aufsetzen des möglichst gestreckten und fixierten linken Beins über die Ferse auf den ganzen Fuß (Gegenstemmen) (Kriterium I37). Dieser sog. Stemschritt, welcher ca. 10–30 Prozent kürzer ist als der Impulsschritt, bremst die Geschwindigkeit des Unterkörpers (Beine und Hüfte) ab, was aufgrund der Massenträgheit zu einer Impulsübertragung auf die oberen, massenmäßig leichteren Körperregionen (Oberkörper, Wurfarm) und somit einer deutlichen Geschwindigkeitszunahme führt (Bauersfeld & Schröter, 2016, S. 599f.). Da speziell das Einnehmen der Wurfauslage als Vorbereitung des Abwurfs eine entscheidende Funktion einnimmt, umfasst die Testvignette zur Wurfauslage jeweils vier Fragen zu den Teilphasen: Druckbeinbewegung, Aufsetzen des Stemschritts und räumlich-zeitliche Ausführung des Stemschritts (siehe Tab. 10).

Tab. 10: Testvignette zur Phase der Wurfauslage

<p>3.1 Kreuzen Sie die Aussagen an, welche bezüglich der Druckbeinbewegung in der Wurfauslage richtig sind. Beachte: Druckbein = hinteres Bein in der Abwurfposition, Stemmbein = vorderes Bein (4 Punkte)</p>	
<p><input checked="" type="checkbox"/> Das Druckbein wird mit einer Beugung im Kniegelenk aufgesetzt.</p> <p><input type="checkbox"/> Die Fußspitze des Druckbeins sollte beim Aufsetzen (entgegen der Wurfrichtung) nach hinten zeigen.</p> <p><input checked="" type="checkbox"/> Nach dem Aufsetzen des Druckbeins wird zuerst die Hüfte nach vorne gedreht und beschleunigt.</p> <p><input checked="" type="checkbox"/> Nach dem Aufsetzen drückt das Druckbein in Wurfrichtung nach vorne gegen das Stemmbein.</p>	
<p>3.2 Kreuzen Sie die Aussagen an, welche bezüglich des Aufsetzens des Stemmbeins richtig sind. (4 Punkte)</p>	
<p><input type="checkbox"/> Die Fußspitze des Stemmbeins zeigt quer zur Wurfrichtung.</p> <p><input type="checkbox"/> Das Stemmbein wird auf dem Fußballen aufgesetzt.</p> <p><input checked="" type="checkbox"/> Das Stemmbein wird gestreckt aufgesetzt.</p> <p><input checked="" type="checkbox"/> Das Stemmbein sollte während des Abwurfs möglichst wenig gebeugt werden.</p>	
<p>3.3 Entscheiden Sie, welche Aussagen bezüglich der räumlich-zeitlichen Ausführung des Stemmschritts richtig sind. (4 Punkte)</p>	
<p><input checked="" type="checkbox"/> Noch bevor das Druckbein aufsetzt, wird das Stemmbein wieder in Richtung Druckbein beschleunigt.</p> <p><input type="checkbox"/> Das Stemmbein wird aufgesetzt, nachdem sich das Druckbein im Abwurf vom Boden löst.</p> <p><input type="checkbox"/> Der zeitliche Abstand zwischen dem Setzen des Druckbeins und dem Setzen des Stemmbeins sollte möglichst groß sein.</p> <p><input type="checkbox"/> Der Stemmschritt ist (räumlich betrachtet) länger als der Impulsschritt.</p>	

Wenn das Stemmbein mit der ganzen Sohle Bodenkontakt erreicht hat, beginnt der Aufbau der Bogenspannung. Durch das Aufrichten des Körpers über das gestreckte und fixierte Stemmbein bei gleichzeitiger Drehung des Oberkörpers um die Körperlängsachse wird die Wurf Schulter in Verbindung mit einer Eindrehbewegung des nach hinten gestreckten Wurfarmellbogens (Kriterium IK39) nach vorne gebracht. Die schnelle Beschleunigung der Wurfhand wird aber noch durch Muskelkräfte verzögert, um eine hohe Vordehnung in der Brust-, Schulter- und Armmuskulatur zu erreichen (Kriterium I38). Ebenso bleibt das Stemmbein weiterhin möglichst gestreckt bzw. wird nur minimal gebeugt, wenn der Lande- druck auf das linke Bein einwirkt, um die Bogenspannung möglichst hoch zu halten und somit die Voraus- setzung für eine dynamische Endbeschleunigung des Wurfgeräts zu schaffen. Der rechte Fuß wird dabei proniert, um die Hüftachse frontal zur Wurfrichtung zu stabilisieren, und behält Kontakt mit dem Boden (Kriterium I36) (Bauersfeld & Schröter, 2016, S. 599). Die Items zur Bewegung der Körperver- wringung sind Teil der Testvignette Abwurfbewegung und sind in Tab. 13 unter der Aufgabe 4.1 dar- gestellt.

Tab. 11: Technikanalysebogen für die Phasen des Stemmschritts und der Bogenspannung (Tidow, 1996a, S. 61)

	HI 31	Free arm	Diagonal sideways pull-back
	HI 32	Throwing arm side	Explosive rotation movement to throwing direction
	I 33	Javelin alignment	Constantly orientated to attitude angle
	I 34	Free arm	Flexed and fixed close to trunk
	I 35	Throwing arm	Slightly flexed / at shoulder height / at the back
	I 36	Support foot	Gliding contact
	I 37	Bracing leg / foot	Minimal yielding / contact with the whole sole
	I 38	Breast / pelvis	Frontal
	IK 39	Throwing arm elbow joint	Rising

Im Moment der höchsten Bogenspannung werden Hüfte und Wurf Schulter in frontaler Position abgebremst und der Ellbogen mit eingedrehtem Wurfarm nach vorne oben beschleunigt (siehe Tab. 12: Kriterien K41 und K40). Die Bogenspannung wird aufgelöst und es kommt zu einer Impulsübertragung vom Rumpf auf den Wurfarm, wobei der Oberkörper durch eine leichte Hüftbeugung nach vorne geneigt wird (Kriterium KL47). Der Ellbogen erreicht in dieser Phase sein Geschwindigkeitsmaximum und wird dann abgebremst. Im letzten Teil der Abwurfbewegung wird die Geschwindigkeit auf Wurfhand und Wurfgerät übertragen, indem die Wurfhand nach vorne oben beschleunigt wird, man spricht von der sog. Unterarmschleuder (Kriterium KL45). Durch das weitere Gestreckthalten bzw. aktive Strecken des Stemmbeins (Kriterium L48) wird die Impulsübertragung verbessert und die Abwurfhöhe vergrößert. Dabei soll es zu einer möglichst geringen seitlichen Abweichung des Körpers von der Wurfrichtungsebene kommen (nach Bauersfeld und Schröter (2016, S.599f.) höchstens 10–12°) und der optimale Abflugwinkel des Wurfgeräts (Kriterium 54) realisiert werden. Der Gegenarm wird während der gesamten Abwurfbewegung seitlich neben der (bei Rechtswerfern linken) Hüfte fixiert (Kriterium L49). Die hieraus resultierenden vier Items zur Ausführung der Wurfbewegung können in Tab. 13 (Aufgabe 4.2) nachvollzogen werden. Das Abfangen der Wurfenergie erfolgt, indem das Stemmbein weiterhin Bodenkontakt hält (Kriterium LM50) und als Hebel für das nach vorne strebende Druckbein dient, welches dann nach dem Fußaufsatz die restliche Bremsarbeit verrichtet und den Körper ins Gleichgewicht bringt (Kriterien MN51 und N52). Da das Abfangen unmittelbar auf die Abwurfbewegung folgt, beinhaltet die Testvignette Abfangen auch drei Items, die sich unmittelbar auf den Moment des Verlassens des Balls von der Wurfhand beziehen (siehe Tab. 13, Aufgabe 5.1).

Tab. 12: *Technikanalysebogen für die Phasen der Schlagbewegung und des Abfangens (Tidow, 1996a)*

<p>K L</p>	<p>Striking position</p> <p>↓</p> <p>Release position</p>	K 40 Throwing arm	Elbow joint is led over shoulder
		K 41 Lower body	"Bent hip" (negative acceleration)
		K 42 Bracing leg	Abutment function with constant angle
		KL 43 Support foot	In gliding contact
		KL 44 Bracing leg	Extending
		KL 45 Throwing arm / hand	Long follow-through / active
		KL 46 Javelin movement	Longitudinal axis constant alignment
		KL 47 Trunk	Marked hip flexion
		L 48 Bracing leg	Completely extended
		L 49 Free arm	Elbow joint at the side of the hip joint
<p>M N</p>	<p>Levering over</p> <p>↓</p> <p>Recovery</p> <p>Javelin flight</p>	LM 50 Bracing foot	Ground contact with ball of the foot / leg working as a lever
		MN51 Support leg	Overtakes bracing leg actively
		N 52 Support leg	Takes over checking and bracing function
		53	Direction of longitudinal axis identical with that of approach
		54	Optimal (for distance and wind) = release angle
55	Release-induced: only rotation about longitudinal axis		

Tab. 13: *Testvignetten zu den Phasen der Abwurfbewegung (4.1 und 4.2) und des Abfangens (5.1)*

<p>4.1 Kreuzen Sie an, welche Aussagen bezüglich der Körperverwindung im Abwurf richtig sind. (4 Punkte)</p> <p><input checked="" type="checkbox"/> Beim Aufsetzen des Stemmbeins wird die Wurfhand mit möglichst gestrecktem Ellbogen zurückgehalten.</p> <p><input checked="" type="checkbox"/> Der Schultergürtel befindet sich zu Beginn der Abwurfbewegung parallel zur Wurfrichtung.</p> <p><input checked="" type="checkbox"/> Die Körperverwindung im Abwurf startet mit dem Vorbringen der wurfseitigen Hüfte.</p> <p><input type="checkbox"/> Die Bogenspannung wird durch aktives Aufdrehen der Wurf Schulter nach hinten erreicht.</p>
<p>4.2 Kreuzen Sie an, welche Aussagen bezüglich der Ausführung der Wurfbewegung richtig sind. (4 Punkte)</p> <p><input checked="" type="checkbox"/> Der Oberkörper knickt im Abwurf seitlich zur Wurfarmgegenseite ab.</p> <p><input checked="" type="checkbox"/> Zu Beginn der Wurfarmbeschleunigung eilt der Ellbogen der Wurfhand voraus.</p> <p><input checked="" type="checkbox"/> Der Gegenarm bleibt beim Abwurf vor dem Körper.</p> <p><input type="checkbox"/> Die Wurfhand befindet sich kurz vor Verlassen des Balles über Kopfhöhe.</p>
<p>5.1 Kreuzen Sie die Aussagen an, welche bezüglich des Abfangens der Wurfbewegung richtig sind. (4 Punkte)</p> <p><input type="checkbox"/> Im Moment des Verlassens des Balls von der Wurfhand befindet sich der Oberkörper in Rücklage.</p> <p><input checked="" type="checkbox"/> Im Moment des Verlassens des Balls von der Wurfhand befindet sich die Schulterachse quer zur Wurfrichtung.</p> <p><input checked="" type="checkbox"/> Das Stemmbein bleibt nach dem Abfangen möglichst gestreckt.</p> <p><input type="checkbox"/> Das Druckbein hat im Moment des Ballabwurfs keinen Kontakt mehr mit dem Boden.</p>

4.2.1.2 Testverfahren für das räumlich-zeitliche Wissen

Wie schon in den eben dargestellten Testvignetten ersichtlich, wurden aufgrund der objektiveren Auswertbarkeit die Teile A (Räumlich-zeitliches Wissen) und B (Biomechanisches Wissen) des Fachwissens-tests mittels eines gebundenen Itemformats als Antwort-Wahl-Verfahren mit Richtig-Falsch-Aufgaben konstruiert. Die Struktur des Prüfungsteils A wird auf der Grundlage der eben erfolgten Sachanalyse mit insgesamt 32 Einzelitems wie folgt festgelegt:

1. Anlauf und Rückführung des Wurfgeräts (4 Items)
2. Impulsschritt (4 Items) -1
3. Wurfauslage mit Stemmschritt (12 Items) -1
4. Abwurfbewegung (8 Items) -1
5. Abfangen (4 Items) -1

Die inhaltliche Auswahl der Antworten wurde aus der Sollwertanalyse (siehe Kapitel 4.2.1.1) abgeleitet. Auf Fragen über den zyklischen Anlaufteil wurde dabei verzichtet, weil dieser im späteren Beurteilungstest keine Rolle spielt. Die Fragenteile 3. *Wurfauslage mit Stemmschritt* und 4. *Abwurfbewegung* wurden aufgrund ihrer Komplexität noch in kleinere inhaltliche Bereiche aufgeteilt:

- 3.1.1 bis 3.1.4: Druckbeinbewegung in der Wurfauslage
- 3.2.1 bis 3.2.4: Aufsetzen des Stemmbeins
- 3.3.1 bis 3.3.4: Räumlich-zeitliche Ausführung des Stemmschritts
- 4.4.1 bis 4.4.4: Körperverwindung im Abwurf
- 4.2.1 bis 4.2.4: Ausführung der Wurfbewegung

Der gesamte Testbogen kann in Tab. 44 des Appendix eingesehen werden. Die Gruppierung von jeweils vier Fragen zu einer Inhaltseinheit wurde bewusst gewählt, um bei der späteren Testanalyse das stren-

gere KPRIM-Verfahren als alternative Auswertungsformen anwenden zu können. Auf diese Weise werden nicht nur die 32 Items einzeln ausgewertet, sondern es wird stärker berücksichtigt, dass mehrere Details über eine Inhaltseinheit gewusst werden müssen, um von einem vertieften Fachwissen ausgehen zu können (siehe Kapitel 5.2.1.3). Im Originaltestbogen ist auch zu erkennen, dass vor der Beantwortung einer Inhaltseinheit Fachbegriffe erklärt werden, um eine Beeinflussung der Testergebnisse durch ein falsches Begriffsverständnis zu vermeiden, wie dies beispielsweise bei Aufgabe 3 zur eindeutigen Unterscheidung von Druckbein und Stemmbein der Fall ist. Die Ergebnisse der Testanalyse werden dann in Kapitel 5.2.1 näher dargestellt und erläutert.

4.2.2 Biomechanik

Meinel et al. (2015) betonen, dass erst grundlegende Kenntnisse und das Wissen über biomechanische Gesetzmäßigkeiten „einen tieferen ‚Analyseblick‘ in die inneren Zusammenhänge und Funktionsmechanismen von Bewegungen und Teilbewegungen“ (S. 382) ermöglichen, um die Bewegungs- und Funktionsstruktur und weitere strukturergänzende Merkmale wie z. B. den Arten der Bewegungskopplung richtig einzuschätzen. Daher ist begründet zu vermuten, dass auch das Wissen über biomechanische Zusammenhänge der Wurfbewegung eine Rolle spielt für die Beurteilungsfähigkeit derselben. Der Rahmentrainingsplan des DLV beginnt das Kapitel Grundlagen des Speerwurfs mit dem Unterkapitel Biomechanische Aspekte, ehe erst später das Technikmodell Speerwurf entwickelt wird (Killing, 2011, S. 137–142). Hieraus wird schon ersichtlich, dass biomechanische Kenntnisse für das Verstehen der technischen Bewegungsausführung fundamental sind und einen wichtigen Bereich des Fachwissens darstellen, der wiederum eine Rolle bei der Bewegungsbeurteilung spielen kann. Die biomechanischen Erkenntnisse über die Wurfbewegung sind aber auch wichtig, wenn es um die Beurteilung der Eigenrealisation der Testpersonen geht, weil sich daraus quantitative Testwerte eruieren lassen, die eine qualitative Bewegungsbeurteilung absichern können (siehe Kapitel 4.4). Darum erfolgt in der nun folgenden Sachanalyse eine sehr detaillierte Erhebung des biomechanischen Wissens, welches aber für die Konstruktion eines Fragebogens über das biomechanische Wissen im Rahmen von Beurteilungsprozessen inhaltlich sehr stark vereinfacht werden muss, weil es sich vielfach schon um sehr vertiefte und spezialisierte Wissensinhalte handelt, die für eine Sportlehrkraft oder einen Leichtathletiktrainer von Kindern und Jugendlichen keine große Praxisrelevanz haben. Die Darstellung der konstruierten Items für Teil B (Biomechanisches Wissen) des Fragebogens erfolgt daher nicht im Rahmen der Sachanalyse, wie sich dies für das räumlich-zeitliche Wissen angeboten hat, sondern wird in Kapitel 4.2.2.2 in entsprechend didaktisch reduzierter Form thematisiert.

4.2.2.1 Sachanalyse des biomechanischen Wissens

Der Anlauf von Top-Speerwerfern ist 25–35 m lang, was den Speerwurf zur Wurfdisziplin mit dem längsten Anlauf macht. Der zyklische Anlauf (bei Topwerfern meist 8–12 Schritte lang) sorgt für den größten Teil der Vorbeschleunigung und wird als Steigerungslauf mit zunehmender Schrittlänge und Schrittfrequenz ausgeführt, weshalb der Begriff zyklisch nur bedingt zutrifft. Adamczewski und Perlt (2003) fanden mittels dynamometrischer Messungen und Analyse synchron aufgezeichneter Videosequenzen bei nationalen Wettkämpfen und leistungsdiagnostischen Erhebungen von über 1000 Würfeln

heraus, dass die Anlaufgeschwindigkeit bei Männern 11–6 m vor der Abwurfline durchschnittlich 6,27 m/s (SD = 0,56) und bei Frauen 10–5 m vor der Abwurfline durchschnittlich 5,74 m/s (SD = 0,36) beträgt. Weltklasseathleten erreichen sogar bis zu 8 m/s und Weltklasseathletinnen bis zu 7 m/s. Trotz dieser schon hohen Geschwindigkeiten erreichen die Werferinnen und Werfer nur 70 Prozent ihrer maximalen Sprintschnelligkeit, weshalb die maximale Sprintleistung damit nicht leistungslimitierend ist (Strüder et al., 2016, S. 671).

Im azyklischen Anlaufteil, dem sogenannten 5er-Rhythmus, soll das Anlauftempo weiter möglichst hoch gehalten werden, denn Murakami et al. (2006, S. 69) berechneten mittels biomechanischer Messungen von 57 Männern mit Wurfweiten zwischen 45,25 m und 87,17 m, dass die Anlaufgeschwindigkeit in der Flugphase des Impulsschrittes der Wert ist, welcher am deutlichsten mit der Wurfweite korreliert ($r = 0,742$, $p < 0,001$). Durch den ziehenden Fußaufsatz vor dem Impulsschritt und seine flache Flugphase wird zwar versucht, weiter zu beschleunigen, was aber selbst Top-Werfern nicht gelingt. Daher kann ein Halten der Geschwindigkeit als Mindestanforderung für den Geschwindigkeitsverlauf des azyklischen Anlaufteils ausreichen (Strüder et al., 2016, S. 670f.), auch wenn Killing (2011, S. 138) eine weitere Geschwindigkeitserhöhung im dritt- und vorletzten Anlaufschritt fordert. Die Horizontalgeschwindigkeit gehe bei Top-Werfern erst im letzten Schritt durch die Impulsübertragung von den Beinen auf Rumpf, Schulter und Wurfarm verloren.

Die letzten fünf Fußaufsätze des azyklischen Anlaufteils sind deutlich weniger aus der Frontalposition gedreht als Hüft- und Schulterachse und zeigen mit der Fußspitze tendenziell stärker in Wurfrichtung (siehe Abb. 34: 20–45° vorge dreht). Der Impulsschritt gehört dabei zu den längsten Schritten des 5er-Rhythmus und stellt die letzte Vorbereitung der Abwurfbewegung dar. Der drittletzte Fußaufsatz sollte dabei nicht zu weit vor der vertikalen Projektion des Körperschwerpunkts erfolgen, da dies entsprechende Bremskräfte bei der Landung verursachen würde. Ebenso bewirkt der flache Abdruck, dass nach dem Aufsetzen des zweitletzten Bodenkontakts (Druckbein) nicht erst eine große Fallgeschwindigkeit kompensiert werden muss, bevor die Abwurfmechanismen starten können. Das doppelte Scheren der Beine (vor dem drittletzten und vorletzten Bodenkontakt) sollte dabei möglichst schnell erfolgen, indem das jeweilige Schwungbein am Stützbein vorbei nach vorne zieht. Noch bevor der vorletzte Bodenkontakt (Druckbeinaufsatz) erfolgt, befindet sich das Stemmbein schon wieder vor dem Druckbein, denn „je weiter der Stemmbeinfuß im Moment des Aufsetzens des Druckbeins vor diesem ist, desto schneller kann das Stemmbein in der Folge zum Boden gesetzt werden“ (Strüder et al., 2016, S. 674). Das Druckbein selbst wird mit einer leichten Beugung im Kniegelenk (auch als Sitzen auf dem Druckbein, weiches Knie oder soft step bezeichnet) etwa auf die vertikale Projektion des Körperschwerpunkts aufgesetzt, wodurch wiederum die Bremskräfte zu Beginn des Druckbeinstützes minimiert werden. Als Orientierung sollte sich die Wurfarm-Gegenschulter beim Setzen des Druckbeins über dem Druckbeinfuß befinden (Killing, 2011, S. 139). Zugleich ermöglicht die Kniebeugung durch eine leichte Absenkung des Körperschwerpunkts erst das Aufsetzen des Stemmbeins deutlich vor dem Körper. Die Fußspitze des Druckbeins zeigt dabei deutlich in Wurfrichtung (weniger als 45° zurückgedreht). Das Druckbein drückt möglichst schnell nach dem Aufsatz aktiv nach vorne, wodurch Druckbeinknie und -hüfte nach vorne gedreht und beschleunigt werden (Strüder et al., 2016, S. 673f.).

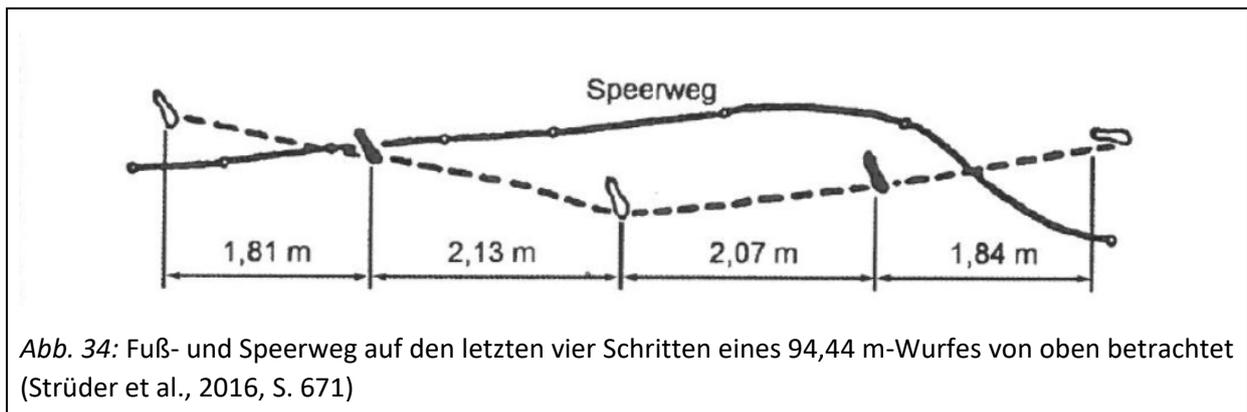


Abb. 34: Fuß- und Speerweg auf den letzten vier Schritten eines 94,44 m-Wurfes von oben betrachtet (Strüder et al., 2016, S. 671)

Die Flugphase des Impulsschrittes dauert ca. 0,20–0,30 s. Je größer die Anlaufgeschwindigkeit, desto länger ist auch der Impulsschritt. Böttcher und Kühl (1998, S. 58) empfehlen beispielsweise für eine 1,78 m große Frau je nach Anlaufgeschwindigkeit folgende Impulsschrittlängen: bei 6,0 m/s 2,00 m, bei 6,5 m/s 2,15 m und bei 7,0 m/s 2,30 m. Der vorhergehende Schritt sollte dann um den Faktor 0,91, der nachfolgende Stemmschritt um den Faktor 0,77 kleiner sein. Killing (2011, S. 138) hält ein Impulsschritt-Stemmschritt-Verhältnis von 3:2 für optimal. Nach Messungen von Adamczewski und Perl (2003) ist der Stemmschritt ungefähr 10–30 Prozent kürzer als der Impulsschritt und somit der kürzeste Schritt des 5er-Rhythmus, wobei sich bei Topwerfern auch große interindividuelle Unterschiede in der Technikausprägung zeigen.³³ Geringe Kraftfähigkeit in der Stemmbeinmuskulatur kann eine Ursache für einen uneffektiven Abwurfvorgang sein, denn wenn die Belastung im Stemmbein zu groß ist und der Kniewinkel des Stemmbeins während des Abwurfs zu sehr nachgibt, erfolgt eine geringere Impulsübertragung der Anlaufgeschwindigkeit auf das Wurfgerät. Je schneller der Anlauf erfolgt, desto mehr Energie kann im Wurfvorgang umgesetzt werden, aber desto besser muss auch die Kraftfähigkeit des Stemmbeins ausgebildet sein (Strüder et al., 2016, S. 674f.). In Abb. 35 werden die hohen Bodenreaktionskräfte sichtbar, die während des Stemmschrittsetzens vom Werfer amortisiert werden müssen (siehe rechte Kraftspitze).

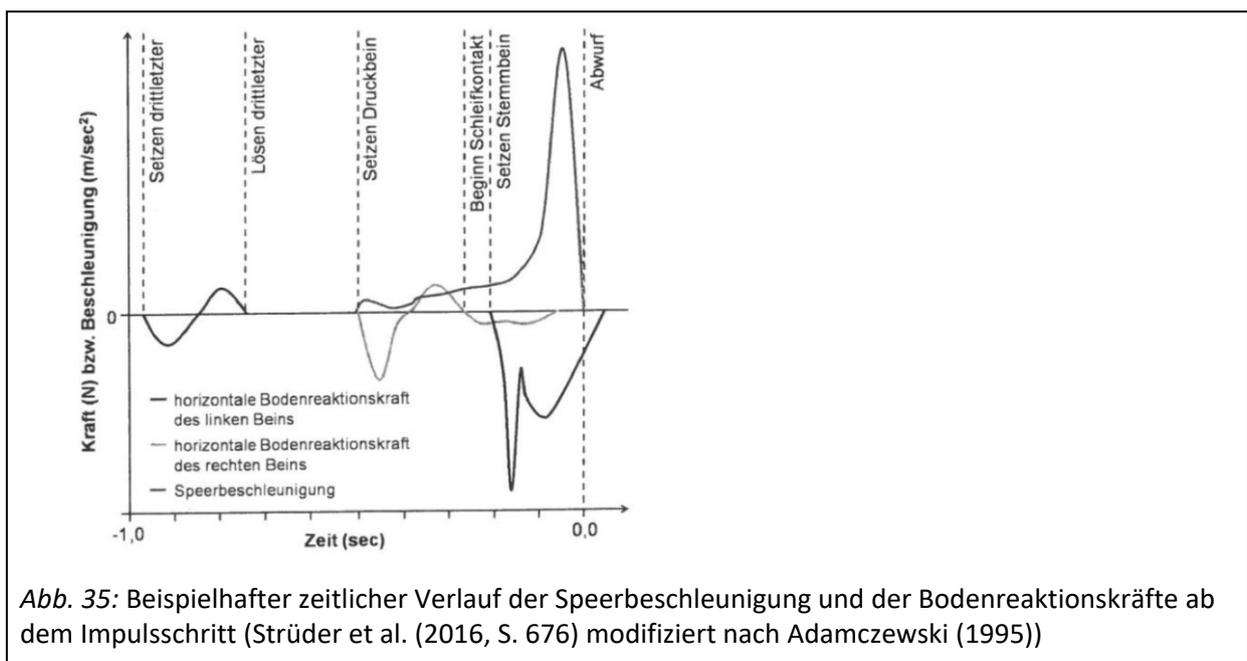


Abb. 35: Beispielhafter zeitlicher Verlauf der Speerbeschleunigung und der Bodenreaktionskräfte ab dem Impulsschritt (Strüder et al. (2016, S. 676) modifiziert nach Adamczewski (1995))

³³ Zelezny führte bei seinem Weltrekordwurf den Stemmschritt sogar geringfügig länger aus als den Impulsschritt (Strüder et al. (2016, S.675)).

Böttcher und Kühl (1996, 80f.) fanden heraus, dass die Größe des Beschleunigungswegs positiv mit der Wurfweite korreliert ($r = 0,73$). Zu diesem Zweck wird auch der Oberkörper mit der Ausführung des Impulsschrittes in eine Rücklage gebracht, welche schon im drittletzten Schritt eingeleitet wurde (siehe Abb. 36). Beim Aufsetzen des Druckbeins ist die Längsachse des Rumpfes um $10\text{--}30^\circ$ von der Vertikalen nach hinten geneigt. Biomechanisch betrachtet wäre aber auf Kosten der Beschleunigungswegverlängerung sogar eine möglichst geringe Rückneigung anzustreben, weil dies wieder zu einer starken exzentrischen Muskelkontraktion zu Beginn des Druckbeinaufsatzes und somit zu einem Geschwindigkeitsverlust bereits vor dem Aufsetzen des Stemmbeins führen würde (Strüder et al., 2016, S. 673f.). Killing (2011, S. 139) spricht daher von einem „Optimum an Rücklage, das einerseits einen langen Beschleunigungsweg der Wurfhand garantiert und andererseits die Bremswirkung nach dem Setzen des Druckbeins gering hält.“

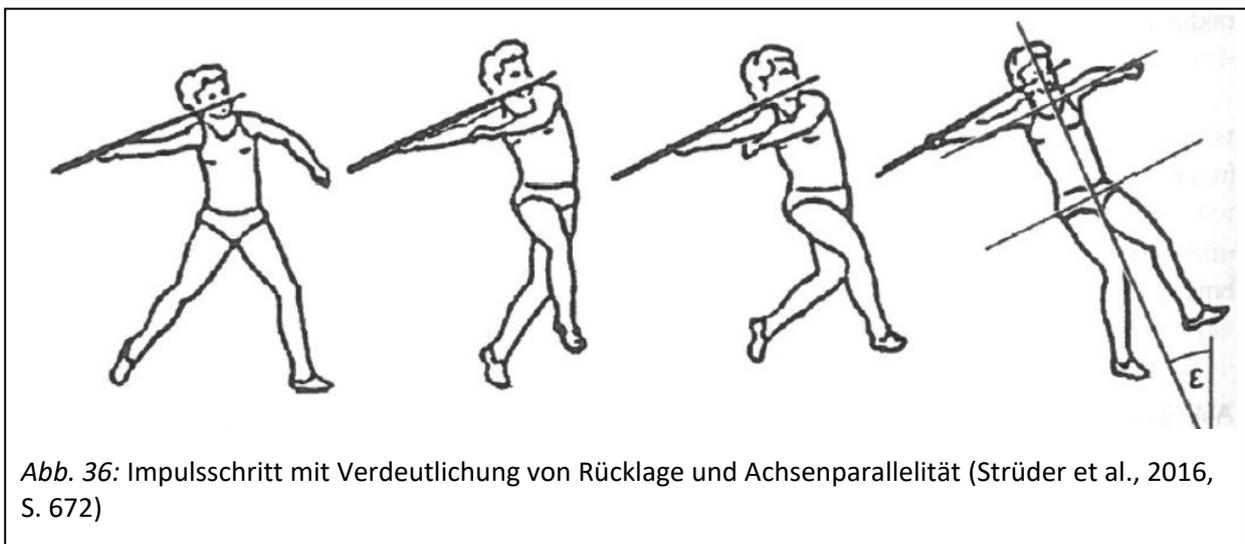


Abb. 36: Impulsschritt mit Verdeutlichung von Rücklage und Achsenparallelität (Strüder et al., 2016, S. 672)

Betrachtet man die Abwurfbewegung, so lässt sich ein Zusammenhang zwischen dem Beschleunigungsweg und der Flugweite feststellen (Böttcher und Kühl (1996) mit $r = 0,73$ und Murakami et al. (2006) nur mit $r = 0,426$). Weltklasseathleten erreichen vom Moment des Druckbeinaufsatzes bis zum Verlassen der Wurfhand einen Beschleunigungsweg von $3,0\text{--}3,5$ m und benötigen dafür eine Zeitspanne von $0,28\text{--}0,38$ s. Das Aufsetzen des Stemmbeins folgt zeitlich nur minimal nach dem Druckbein (bei Topwerfern $0,14\text{--}0,26$ s). Weil das Druckbein meist noch ganzsohlig Bodenkontakt hat oder nach seiner Fußgelenkstreckung auf der Fußoberseite schleift, wenn das Stemmbein aufsetzt, ist der Stemschritt letztlich ein Gehschritt. Die eigentliche Abbremsbewegung der Anlaufgeschwindigkeit erfolgt vom Stemmbeinaufsatz bis zum Moment des Verlassens des Wurfgeräts aus der Hand. Bei Topwerfern erfolgt innerhalb von $0,11\text{--}0,14$ s eine Geschwindigkeitsabnahme um $50\text{--}70$ Prozent auf $1,5\text{--}3,5$ m/s (Strüder et al., 2016, S. 675). Dieses starke Abbremsen gelingt nur, wenn das Stemmbein möglichst wenig gebeugt wird. Killing (2011, S. 141) fordert einen Kniewinkel im Stemmbein zum Zeitpunkt des Setzens von 160° und mehr, wobei Adamczewski und Perl (2003) selbst bei ausgebildeten Speerwerferinnen und Speerwerfern für den minimalen Kniewinkel im Stemmbein lediglich Werte zwischen $146\text{--}153^\circ$ feststellten. Die aktive Stemmbeinarbeit begünstigt auch die Oberkörpervorlage im Moment des Abwurfs, welche positiv mit der Wurfweite korreliert ($r < 0,463$, $p < 0,01$) und als Indiz für die Bedeutung einer hohen Vorwärtsrotationsgeschwindigkeit des Oberkörpers gedeutet werden kann (Murakami et al., 2006).

Betrachtet man die Kopplung der Teilbewegungen im Abwurfvorgang, so stellt man fest, dass die Geschwindigkeitsmaxima einzelner Körperteile kurz nacheinander realisiert werden. Bei den von Lehmann (2005) durchgeführten Analysen bei den Deutschen Meisterschaften von 2004 zeigte sich, dass die wurfseitige Hüfte kurz vor dem Aufsetzen des Stemmbeins (0,08–0,12 s) ihr Geschwindigkeitsmaximum erreicht, gefolgt von der Wurfarmschulter fast zeitgleich mit dem Stemmbeinaufsatz und der maximalen Ellenbogengeschwindigkeit ca. 0,05 s später. Das Abbremsen des Ellbogens, die Beschleunigung der Wurfhand durch die Ellbogenstreckung und die abschließende Handgelenksstreckung erfolgt unmittelbar danach (0,01–0,03 s später). Campos, Brizuela und Ramón (2004) ermittelten für sieben Weltspitzenspeerwerfer, dass die wurfseitige Hüfte 0,13 s, die wurfseitige Schulter 0,09 s und der wurfseitige Ellbogen 0,06 s vor dem Abwurf den jeweiligen Geschwindigkeitshöhepunkt erreicht. In Abb. 37 ist nochmals graphisch dargestellt, wie nach der konstanten Anfangsbeschleunigung durch den Anlauf die jeweils energieabgebende größere Einheit (z. B. die Hüfte) in Stemmschritt und Abwurf deutlich an Geschwindigkeit zugunsten der energieaufnehmenden kleineren Einheit (z. B. Schulter) verliert (Killing, 2011, S. 137f.).

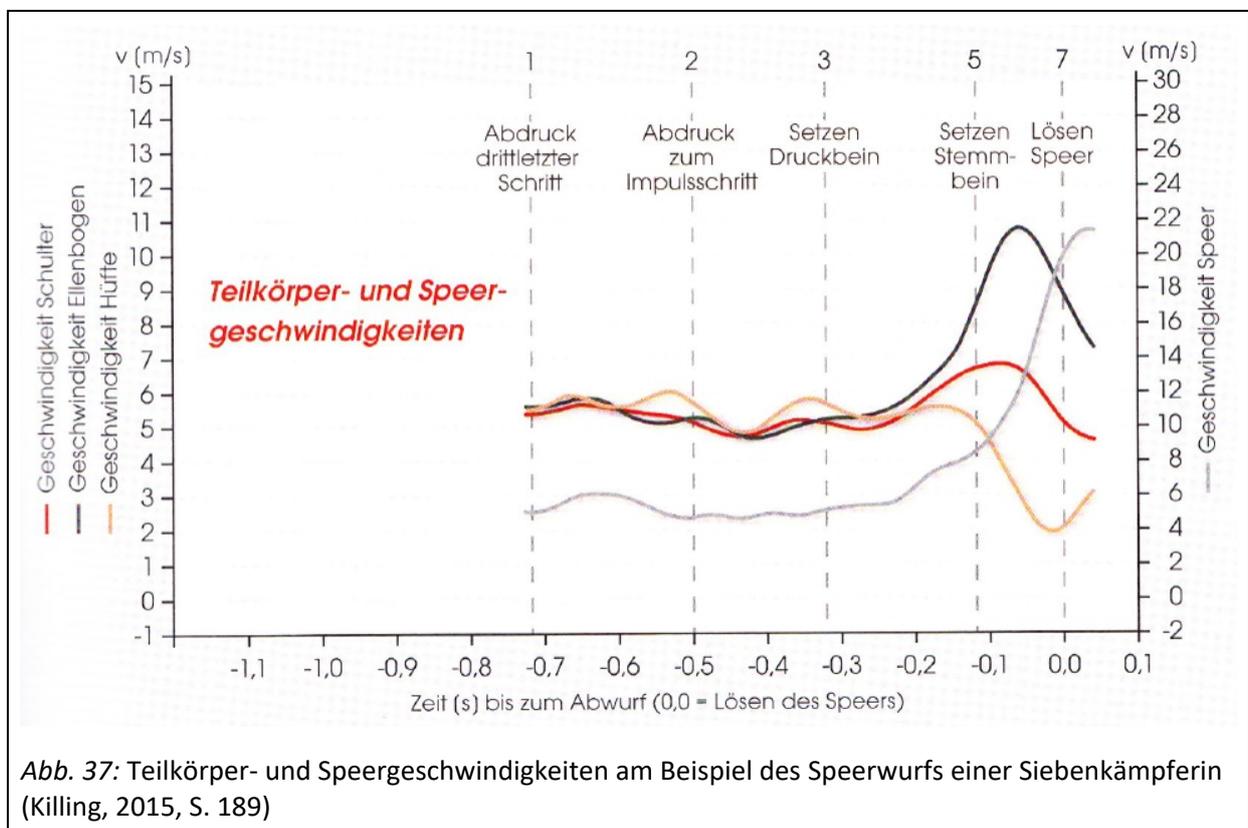


Abb. 37: Teilkörper- und Speergeschwindigkeiten am Beispiel des Speerwurfs einer Siebenkämpferin (Killing, 2015, S. 189)

Die Wurfverzögerung kann auch am Ellbogenwinkel im Moment des Stemmbeinaufsatzes gemessen werden. Leistungsstärkeren Athleten gelingt es besser, den Wurfarm möglichst lange gestreckt hinter dem Körper zu halten bzw. einen größeren Ellbogenwinkel beim Stemmbeinaufsatz zu realisieren. Adamczewski und Perlt (2003) konnten für Speerwurfspezialisten Ellbogenwinkel von 140° beim Stemmbeinaufsatz feststellen, während Mehrkämpfer nur 132° erreichten. Messungen des Geschwindigkeitsverlaufs des Wurfgeräts während der Abwurfbewegung zeigen, dass das Wurfgerät in der zeitlichen Mitte zwischen Stemmbeinaufsatz und Verlassen des Wurfgeräts von der Hand noch nicht 40 Prozent seiner Endgeschwindigkeit erreicht haben sollte. Top-Werfer schaffen es demnach, 60 Prozent der kinetischen Energie erst in den letzten 0,05 s des Abwurfs auf das Wurfgerät zu übertragen (Strüder

et al., 2016, S. 677f.). Dabei muss die Theorie der Unterarmschleuder um zwei Aspekte erweitert werden. Böttcher und Kühl (1996) stellen (mit $r = 0,82$) fest, dass bei Top-Werfern weite Würfe mit einer geringeren Ellbogenbeugung im Abwurf ausgeführt werden, woraus sie folgern, dass der Winkel im Ellbogen optimaler Weise rechtwinklig sein soll. Murakami et al. (2006) bestätigen die Ergebnisse von Campos et al. (2004), dass die Ellbogenstreckung am Ende des Abwurfs mit $60\text{--}70^\circ$ nur relativ gering ausfällt und der Ellbogenwinkel beim Verlassen des Wurfgeräts ca. $150\text{--}160^\circ$ beträgt. Strüder et al. (2016, S. 679) fassen diese Ergebnisse daher wie folgt zusammen:

Murakami et al. (2006) schlagen daher vor, die Unterarmschleuder mehr als schnelle Innenrotation im Schultergelenk mit relativ konstantem Ellbogenwinkel zu betrachten. Bei einer Rotation, deren Drehachse durch den Oberarm verläuft, wird durch einen Ellenbogenwinkel von 90° der größte Radius (Hebel) zwischen Drehachse und Speer hergestellt, um die Rotation bestmöglich in Speergeschwindigkeit umzusetzen.

Ein weiteres Detail der Abwurfbewegung stellt die Wurfseitenfreiheit dar (siehe Abb. 38). Gute Werfer neigen ihren Oberkörper im Abwurf zur Stemmbeinseite, weil in dieser Seitneigung (mit kleiner Abduktion im Schultergelenk) größtmögliche beschleunigungswirksame Kräfte (z. B. das Nach-vorne-Führen und die Innenrotation des Schultergelenks durch den Pectoralis) erzeugt werden können. Murakami et al. (2006) folgern aus ihren Messungen, dass mit einer mittleren Seitneigung bei einem Abduktionswinkel von ca. 90° neben einer effektiven Muskularbeit auch ein größtmöglicher Abwurfwinkel und eine ausreichende Abflughöhe verwirklicht werden können. Ebenfalls vorteilhaft an einer Seitneigung ist die Tatsache, dass dadurch der Körperschwerpunkt des Werfers leichter hinter das Stemmbein gebracht und die Stemmwirkung somit verbessert werden kann (Strüder et al., 2016, S. 679f.).

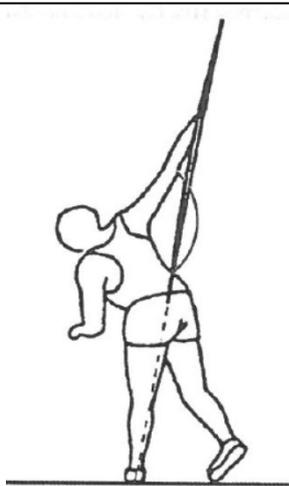


Abb. 38: Speerwerfer mit ausgeprägter Wurfseitenfreiheit aus der Rückperspektive (Strüder et al., 2016, S. 679)

Räumlich betrachtet muss der Abwurf so gestaltet werden, dass der Werfer die Abwurfslinie nicht übertritt, aber trotzdem eine möglichst explosiv vorwärts gerichtete Abwurfbewegung ausführen kann. Wenn der Abwurf in der entsprechenden Dynamik ausgeführt wird, ist ein Abfangen des Körpers mit einem zusätzlichen Schritt nach vorne unabdingbar. Hierzu müssen schon im Moment des Abwurfs einige Bewegungsmerkmale verwirklicht sein: der Körperschwerpunkt sollte den Aufsatzzpunkt des Stemmbeins noch nicht überschritten haben und das Stemmbein sollte möglichst gestreckt gehalten

werden, auch wenn das Wurfgerät die Hand bereits verlassen hat. Mit einem weiten Abfangschritt, der deutlich vor der vertikalen Projektion des Körperschwerpunkts aufgesetzt wird, muss die übrige Horizontalgeschwindigkeit abgebremst werden (Strüder et al., 2016, S. 678–680).

Bei den meisten biomechanischen Analysen werden die Parameter Abfluggeschwindigkeit, Abflugwinkel und Abflughöhe gemessen. Die Abfluggeschwindigkeit (bei Top-Männern 28–32 m/s, bei Top-Frauen 24–28 m/s) ist dabei der entscheidende Faktor und bestimmt zu ca. 80 Prozent die Wurfweite ($r = 0,9$). Die Abflugwinkel, gemessen durch die Bahn der Speerwicklung im Moment des Abwurfs und 0,04 s danach, variieren bei guten Speerwerfern sehr stark, im WM-Finale 1999 wurden Werte zwischen 27,7–40,1° gemessen, wobei die Messwerte keine geschlechts-, alters- und leistungsbezogenen Tendenzen aufweisen (Adamczewski & Perlt, 2003). Es muss auch darauf hingewiesen werden, dass der optimale Abflugwinkel von den Windverhältnissen abhängt, denn je mehr Gegenwind herrscht, desto flacher muss der Speer auch abgeworfen werden. Theoretisch wäre ein Abflugwinkel zwischen 40–45° für die Erzielung einer großen Flugweite des Speeres optimal, jedoch kann bei einem derart großen Abwurfwinkel der Oberkörper und Arm nicht so gut nach vorne beschleunigt werden, wodurch Abfluggeschwindigkeit eingebüßt wird. Daher muss der Werfer versuchen, nach individuellen Technikmerkmalen einen möglichst großen Abflugwinkel zu realisieren, ohne dabei unverhältnismäßig viel Abwurfgeschwindigkeit zu verlieren. Killing (2011, S. 140) empfiehlt bei geringem Wind einen Winkelbereich zwischen 34–38°. Die Abwurfhöhe ist zwar auch bei Top-Athleten sehr unterschiedlich (zwischen 1,7–2,1 m), zeigt aber keinen statistischen Zusammenhang mit der Wurfweite (Strüder et al., 2016, S. 680). Auf Angriffs- und Verkantungswinkel des Speers wird nicht mehr näher eingegangen, weil diese biomechanischen Parameter für die Wurfbewegung eines Balles keine Rolle spielen.

4.2.2.2 Testverfahren für das biomechanische Wissen

Als biomechanisch werden vor allem Aussagen gewertet, die über die rein räumlich-zeitliche Dimension einer Bewegung hinausgehen und eine mechanische oder physikalische Sichtweise auf die Bewegung voraussetzen. Am offensichtlichsten werden biomechanische Erkenntnisse immer dann, wenn quantitativ messbare Werte erhoben werden (wie z. B. Winkelangaben, Geschwindigkeiten, Schrittlängen). Aber auch Begriffe wie Körperschwerpunkt, Beschleunigungsweg oder Bodenreaktionskräfte deuten darauf hin, dass die damit zusammenhängenden Aussagen ohne biomechanisches Wissen nicht verstehbar sind. Darum wurde auch bei der Itemkonstruktion darauf geachtet, dass die zu bewertenden Aussagen sich auf biomechanische Begriffe und Zusammenhänge beziehen.

Aus den doch sehr umfangreichen und detaillierten biomechanischen Informationen der Sachanalyse (siehe Kapitel 4.2.2.1) über die leichtathletische Schlagwurfbewegung (konkretisiert am Beispiel des Speerwurfs) wurden nach einer starken inhaltlichen Reduktion 16 Items entwickelt, welche in vier Testvignetten zusammengefasst wurden: 1. Azyklischer Anlauf (Vignette 6.1), 2. Wurfarmbewegung (Vignette 6.2), 3. Körperverwindung und Bogenspannung (Vignette 6.3) und 4. Abfangen (Vignette 6.4) (siehe Tab. 14). Ausschlaggebend für die inhaltliche Selektion war die Frage, ob der entsprechende biomechanische Wissensinhalt relevant sein kann für die Beurteilung der sportlichen Bewegung. Das Wissen, dass der Impulsschritt nicht zum Abbremsen, sondern zum Beschleunigen der Horizontalgeschwindigkeit dient (siehe Tab. 14, Item 6.1.1), ist beispielsweise entscheidend, um das Bewegungstempo des Anlaufs richtig einschätzen zu können. Wenn Beurteilende wissen, dass der Stemmbeinfuß umso schneller in der Wurfauslage gesetzt werden kann, je weiter er sich im Moment des Aufsetzens

des Druckbeins vor diesem befindet (siehe Tab. 14, Item 6.1.3), so kann möglicherweise der Bewegungsrhythmus der letzten drei Bodenkontakte besser beurteilt werden oder eine zielführende Bewegungskorrektur erfolgen. Wenn es um die Beurteilung der sehr schnell ablaufenden Wurfbewegung, Körperverwindung oder Bogenspannung geht, so kann eine biomechanisch begründete Bewegungsvorstellung helfen, die kompliziert zu beobachtenden Merkmale wie Bewegungsumfang oder Bewegungskopplung angemessen zu bewerten. Vor allem auf das Wissen über die Kopplung der Teilimpulse einzelner Körperteile durch Beschleunigen und Verzögern wurde bei der Itemkonstruktion Wert gelegt, weil diese Bewegungsfertigkeit ein zentrales Element einer gut ausgeführten Schlagwurfbewegung darstellt (siehe Tab. 14, Items 6.2.1, 6.2.3, 6.2.4, 6.3.1, 6.3.3 und 6.3.4). Bei der letzten Testvignette über das Abfangen sind die inhaltlichen Aussagen vor allem auf Teilkörperbewegungen gerichtet, die hilfreich sind, um die Horizontalgeschwindigkeit kontrolliert zu reduzieren. Das Wissen über die Funktion der richtigen Körperposition beim Abfangen kann helfen, das Merkmal Bewegungspräzision im Bezug auf ein mögliches Übertreten der Abwurfmarkierung einzuschätzen. Als Erhebungsmethode wird aufgrund der objektiveren Auswertung wieder die gebundene Antwortform gewählt, indem vorgegebene biomechanische Behauptungen als richtig oder falsch eingeschätzt werden müssen (siehe auch den Gesamtfragebogen im Appendix, Tab. 44).

Tab. 14: *Testvignetten zum biomechanischen Wissen*

<p>6.1 Kreuzen Sie die Aussagen an, welche bezüglich der letzten drei Anlaufschritte richtig sind. (4 Punkte)</p> <p><input type="checkbox"/> Mit dem Impulsschritt beginnt das Abbremsen der Horizontalgeschwindigkeit.</p> <p><input checked="" type="checkbox"/> Das Druckbein wird im vorletzten Bodenkontakt ungefähr auf die vertikale Projektion des Körperschwerpunktes auf dem Boden aufgesetzt.</p> <p><input checked="" type="checkbox"/> Je weiter der Stemmbeinfuß im Moment des Aufsetzens des Druckbeins vor diesem ist, desto schneller kann das Stemmbein in der Folge zum Boden gesetzt werden.</p> <p><input checked="" type="checkbox"/> Der Gegenarm unterstützt durch sein langes Verharren in Wurfrichtung das späte Drehen der Schulterachse in die Frontalposition.</p>
<p>6.2 Kreuzen Sie die Aussagen an, welche bezüglich der Bewegung des Wurfarms richtig sind. (4 Punkte)</p> <p><input checked="" type="checkbox"/> Am Ende der Wurfbewegung wird der Ellbogen kurz vor seiner Streckung abgebremst.</p> <p><input checked="" type="checkbox"/> Die Handgelenksstreckung übt den letzten Beschleunigungsimpuls auf das Wurfgerät aus.</p> <p><input type="checkbox"/> Aus der Wurfauslage erfolgt das Beschleunigen und Abbremsen der einzelnen Wurfarmgelenke zeitlich betrachtet von distal (fern vom Körperzentrum) nach proximal (nah am Körperzentrum).</p> <p><input type="checkbox"/> Der Ellbogen des Wurfarms wird beim Abwurf zeitgleich mit der Wurfhand nach vorne gebracht.</p>
<p>6.3 Kreuzen Sie die Aussagen an, welche bezüglich der Körperverwindung und Bogenspannung richtig sind. (4 Punkte)</p> <p><input checked="" type="checkbox"/> Kurz vor dem Aufsetzen des Stemmbeins erreicht die wurfseitige Hüfte ihr Geschwindigkeitsmaximum.</p> <p><input checked="" type="checkbox"/> Die entgegengesetzt wirkende Druck- und Stemmbeinarbeit führt zur Verwindung zwischen Hüft- und Schulterachse.</p> <p><input checked="" type="checkbox"/> Mit dem Vorbringen der Wurfarmschulter bei gleichzeitiger Verzögerung des Wurfarms ist der Aufbau der Bogenspannung abgeschlossen.</p> <p><input checked="" type="checkbox"/> Der Wurfarm sollte solange verzögert möglichst gestreckt hinter dem Körper bleiben, bis die wurfarmseitige Schulter nach der Beschleunigung wieder abgebremst wird.</p>
<p>6.4 Kreuzen Sie die Aussagen an, welche bezüglich des Abfangens der Wurfbewegung richtig sind. (4 Punkte)</p> <p><input type="checkbox"/> Im Moment des Abwurfs sollte der Körperschwerpunkt den Aufsatzpunkt des Stemmbeins bereits überschritten haben.</p> <p><input checked="" type="checkbox"/> Die Rückführung des Gegenarms an den Körper beim Abwurf unterstützt die Längsachsenrotation des Oberkörpers.</p> <p><input checked="" type="checkbox"/> Mit einem Schritt über das Stemmbein hinweg kann die noch vorhandene Horizontalgeschwindigkeit aus dem Abwurf abgefangen werden.</p> <p><input checked="" type="checkbox"/> Das Stemmbein bleibt beim Abfangen möglichst lang hinter dem Körper zurück, um den Körperschwerpunkt möglichst weit hinten zu halten.</p>

4.2.3 Fehlerbilder

Blischke, Schumacher und Daug (1993) konnten bereits nachweisen, dass eine fundierte merkmalsbezogene und sollwertorientierte Wissensbasis über eine sportliche Bewegung zu einem hohen Grad der Fehleridentifikation führt. „Inwieweit auch eine *fehlerbezogene* Wissensbasis die Fehleridentifikation beeinflusst [sic], ist bisher empirisch nicht geprüft“ (Marschall et al., 1996, S. 196). Aus diesem Grund wird das Fehlerwissen der Probandinnen und Probanden in dieser Studie erhoben. Neumaier (1988, S. 335) fordert bereits vor 30 Jahren, dass „für die Schulung der Fähigkeit zur detaillierten Bewegungsausführung [...] nicht nur eine detaillierte Vorstellung von der richtigen Bewegungsausführung geschaffen werden muß [sic], sondern in ihr auch Fehlerbilder Berücksichtigung finden sollten.“ Damit man also überhaupt Fehler identifizieren kann, muss ein Beobachter einen mehr oder weniger detaillierten und präzise beschriebenen Sollwert mental abrufbar im Gedächtnis haben (siehe Kapitel 4.2.1.1). Dieser Sollwert ist in der Leichtathletik (abgesehen von kleineren Freiheitsgraden in der individuellen technischen Ausführung einer Bewegung) relativ stark determiniert und somit der Toleranzbereich für Abweichungen von der idealen Bewegungsausführung relativ gering. Trotzdem finden sich in der Fachliteratur zahlreiche Fehlerbilder, die so stark von der idealen Bewegungsausführung abweichen, dass sie schon bei ungeübten Werferinnen und Werfern korrigiert werden sollen, um spätere Lernblockaden zu vermeiden. Wenn beispielsweise ein Athlet das falsche Stemmbein beim Abwurf nach vorne setzt, so hat er auch keine Möglichkeit, die Hüfte bei der Rückführung des Wurfarms mit nach hinten zu nehmen, wodurch wiederum die Bogenspannung nur ungenügend aufgebaut werden kann. Wenn einem Beurteiler die häufigsten Fehlerbilder bekannt sind, so ist anzunehmen, dass er diese bei der Bewegungsbeobachtung auch leichter erkennt. Es wird demnach von der Annahme ausgegangen, dass die eben zitierte fehlerbezogene Wissensbasis über die Fehlerbeschreibungen in der leichtathletischen Standardliteratur operationalisiert werden kann.

4.2.3.1 Sachanalyse des Fehlerwissens

Die Sachanalyse des Fehlerwissens erfolgt erneut über eine umfassende Literaturrecherche. Dabei werden folgende wichtige Standardwerke (siehe Tabelle 15) mit Aussagen über Fehlerbeschreibungen der Disziplinen Speerwurf und Ballwurf synoptisch zusammengearbeitet.

Tabelle 15: *Literatur über Fehlerbeschreibungen im Speerwurf und Ballwurf*

Autoren	Überschriften
Strüder et al. (2016, S. 702–705)	Fehler – Ursachen – Korrekturen
Bauersfeld & Schröter (2016, S. 615f.)	Wesentliche Fehler und Korrekturmöglichkeiten
Kunz (2001, S. 22f.)	Fehlerkorrektur Speerwerfen
Haberkorn & Plaß (1992, S. 155f.)	Fehlerbilder – Ursachen – Korrekturmaßnahmen Ballwurf
Haberkorn & Plaß (1992, S. 173)	Fehlerbilder – Ursachen – Korrekturmaßnahmen Speerwurf

Killing (2015, S. 130)	Technische Fehler und Korrekturmöglichkeiten beim Speerwurf
Killing (2011, S. 153–158)	Technikfehler, Ursachen und Korrekturen beim Speerwurf
Stahl (1988)	Häufige Fehler und ihre Korrektur
Weiler & Lindemann (2010)	Fehler, Ursachen und Korrektur

Aus diesen Fehlerbeschreibungen wurde ein Erwartungshorizont konstruiert, in welchem auch die genannten Ursachen einer fehlerhaften Bewegungsausführung aufgeführt werden. Inhaltlich ähnliche Aussagen wurden zu gemeinsamen Fehlerbildern zusammengefasst, so dass für jede Bewegungsphase des Schlagwurfs ein Fehlerkatalog entstand, anhand dessen die Auswertung der Antworten der Probandinnen und Probanden erfolgte. Die Miteinbeziehung der Ursachen erweitert den Bewertungsspielraum bei der Auswertung der Fragebögen und wird der Tatsache gerecht, dass die Kenntnis der Fehlerursache meist didaktisch wertvoller ist als die Identifizierung des Fehlerbildes selbst. Die Erwartungshorizonte sind demnach unterteilt in die Phasen azyklischer Anlauf (siehe Appendix Tab. 46, 15 Fehlerbilder), Rückführung (Appendix Tab. 47, 5 Fehlerbilder), Abwurf (Appendix Tab. 48, 38 Fehlerbilder) und Abfangen (Appendix Tab. 49, 2 Fehlerbilder). Ein auf die Nennung der Fehlerbilder ohne Ursachenbeschreibung reduzierte Fehlerkatalog ist exemplarisch für die Phase des azyklischen Anlaufs in Tab. 16 dargestellt.

Tab. 16: Erwartungshorizont für das Fehlerwissen der Phase Anlauf azyklisch (reduziert auf die Nennung der Fehlerbilder ohne Ursachenbeschreibungen)

1. Fehlerbilder Anlauf azyklisch
A) Oberkörperposition/Körperlage
1. Ungenügende Körperrücklage
2. Zu geringe Körper-/Oberkörperrücklage nach Impulsschritt (bzw. beim Aufsetzen des Druckbeins)
3. Körpervorlage
4. Zu große Oberkörperrücklage beim Aufsetzen des Druckbeins
5. Keine Achsenparallelität von Schulter-, Hüft- und Speerachse beim Setzen des Druckbeins
B) Wurfarm
6. Vorzeitiges Anziehen bzw. Einbeugen des Wurfarmes im Anlauf / nach der Rückführung
C) Impulsschritt
7. Zu hoher, sprunghafter Impulsschritt (einhergehend mit Horizontalgeschwindigkeitsverlusten beim Abdruck zum Impulsschritt und Aufsetzen des Druckbeins)
8. Impulsschritt zu kurz / ohne Doppelwechsel der Beine
9. Falscher Abwurfrhythmus

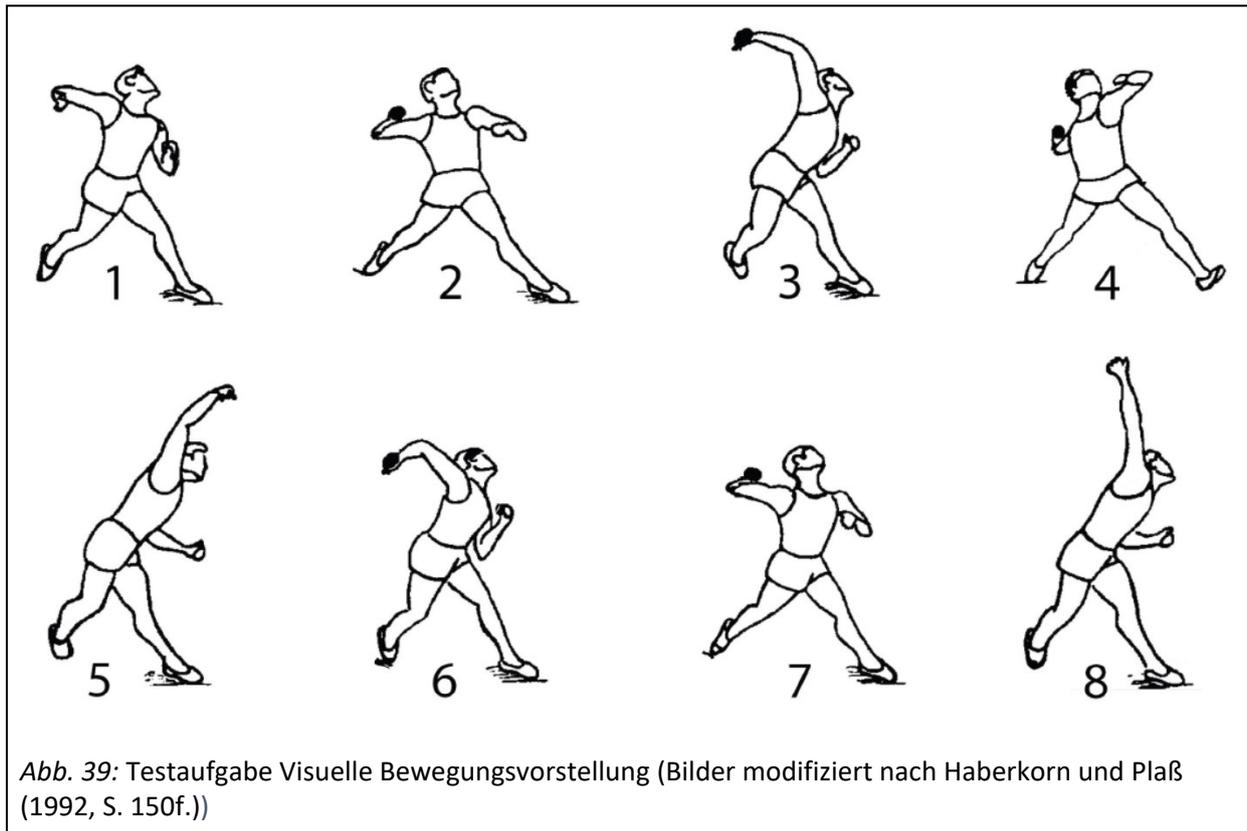
10. Impulsschritt führt nicht in die Wurfauslage, das Gewicht liegt zwischen den Beinen und nicht auf dem rechten Bein, das linke Bein kann nicht als Stemmbein eingesetzt werden
11. Schlechte Schrittgestaltung
D) Geschwindigkeitsabfall
12. Abfall der Geschwindigkeit zu Beginn des Fünf-Schritt-Rhythmus bzw. mangelnde oder fehlende Steigerung der Geschwindigkeit
E) Impulshopser
13. Impulshopser/Nachstellschritt, Nachstellhupf
F) Lauf- und Fußkoordination
14. koordinative Probleme beim seitlichen Laufen mit zurückgenommenem Speer (z. B. Schulterachse in Laufrichtung, Füße quer zur Laufrichtung, eher Seitgalopp statt Kreuzschritt, ungünstige Blickrichtung o.Ä.)
15. Füße zeigen nicht in Laufrichtung / Aufsetzen des rechten Fußes quer zur Wurfrichtung

4.2.3.2 Testverfahren für das Fehlerwissen

Beim Testverfahren Fehlerwissen wurde ein freies Antwortformat gewählt, um die Probandinnen und Probanden nicht durch vorgegebene Antworten in ihrer individuellen Auswahl von wichtigen Bewegungsabweichungen zu beeinflussen (siehe Tab. 44 im Appendix, Vignetten 7.1, 7.2 und 7.3). Die Testpersonen mussten typische Fehler stichpunktartig beschreiben, mit denen bei der Ausführung eines Schlagwurfs durch Anfängerinnen und Anfänger zu rechnen ist. Unterteilt war die Aufgabe in die drei Bewegungsphasen Anlauf (5 Fehlerbilder), Abwurf (5 Fehlerbilder) und Abfangen (2 Fehlerbilder). Es wurde explizit in der Aufgabenbeschreibung darauf hingewiesen, dass das Einnehmen der Wurfauslage nicht mehr zum Anlauf zählt, sondern zur Abwurfphase. Für Anlauf und Abwurf wurden jeweils fünf Fehlerbilder verlangt, für das Abfangen zwei.

4.2.4 Visuelle Bewegungsvorstellung

Bei der letzten Testaufgabe soll festgestellt werden, ob die Probandinnen und Probanden neben den verbal präsentierten Fragen auch eine auf der visuellen Bewegungsvorstellung basierende Aufgabe beantworten können (siehe Tab. 44 im Appendix, Vignette 8.1). Ähnlich dem Bildkartenauswahltest von Daug's et al. (1989) (siehe Kapitel 2.3.1.4.5) werden einzelne Körperpositionen der Schlagwurfbewegung aus ihrer richtigen, dem chronologischen Bewegungsverlauf entsprechende Reihenfolge gebracht und durcheinander angeordnet. Die Testpersonen müssen die mit Zahlen kodierte Bilder neu sortieren, so dass die Bewegungsabfolge der Einzelbilder wieder stimmt (siehe Abb. 39). Über die entsprechende Zahlenkombination wird die Richtigkeit des Ergebnisses kontrolliert bzw. über fehlerhafte Positionen eine Schwäche in der visuellen Bewegungsvorstellung festgestellt.



4.2.5 Gütekriterien

Die Überprüfung der **Gütekriterien** eines Testverfahrens stellt einen wichtigen Bestandteil der **Testanalyse** (siehe Schritt 11 des Testalgorithmus nach Bös (2017a) in Abb. 31) dar. Anhand der drei Hauptgütekriterien Objektivität (siehe Kapitel 4.2.5.1), Reliabilität (siehe Kapitel 4.2.5.2) und Validität (siehe Kapitel 4.2.5.3), deren Prüfung und Einhaltung bei der Testkonstruktion wissenschaftsübergreifend als unverzichtbar angesehen wird, soll der wissenschaftliche Anspruch des eben beschriebenen Fachwissenstests begründet werden.

4.2.5.1 Objektivität

Die 59 Testpersonen führen den Fachwissenstest in drei Gruppen durch, wobei die ersten beiden Gruppen den Test zeitgleich in unterschiedlichen Räumen absolvierten. Die **Durchführungsobjektivität** wird durch eine genaue Durchführungsanleitung für die Testleiter gewährleistet, in welcher alle kritischen Fragen und Unklarheiten besprochen werden. Vor allem das genaue Vorgehen zu Beginn des Tests (z. B. Vorinformationen über Ziel und Inhalt des Tests, Anmerkungen zum Ausfüllen des Fragebogens) sowie während der Beantwortung (Austeilen und Einsammeln der drei Antwortbögen, Reaktion auf eventuell auftretende Fragen und Unklarheiten) ist darin festgelegt. Eine hohe **Auswertungsobjektivität** wird zum einen durch die Verwendung der Testsoftware KLAUS, mit welcher der Fachwissenstest auch konstruiert wird, gewährleistet. Das Einlesen der Antwortbögen erfolgt mittels Scanner, wodurch

menschliche Flüchtigkeitsfehler bei der Codierung der Antworten vermieden werden. Die eingescannten Bögen werden durch die Software kontrolliert und alle Antworten, welche kein eindeutiges Ankreuzen erkennen lassen, dem Auswerter vorgelegt, welcher die Antworten nach Sichtkorrektur und eventuellem Vergleich mit dem Originaldokument manuell als richtig oder falsch bestätigt.

Die Teile A und B können durch das geschlossene Antwortformat ausschließlich softwaremäßig ausgewertet werden. Die freien Antworten von Teil C müssen manuell geprüft und mit dem Codebook Fehlerwissen (siehe Tab. 45 im Appendix) verglichen werden, ehe der erreichte Punktwert durch den Auswerter auf dem Originalbogen angekreuzt werden kann. Dazu werden alle Antworten aus Teil C manuell transkribiert und in Tabellenform gebracht, so dass alle Antworten der einzelnen Probanden in einer Tabellenspalte untereinander angeordnet sind und im Vergleich zueinander möglichst kohärent ausgewertet werden können. Es wird festgelegt, dass eine Antwort mit zwei Bewertungseinheiten (BE) gewertet wird, wenn sie mit einer Antwortmöglichkeit des Codebooks inhaltlich übereinstimmt. Antworten, die sich in der Fachliteratur nur in den Ursachenbeschreibungen eines Fehlers finden lassen (ohne jedoch den Fehler selbst zu nennen), werden mit 1 BE bepunktet. Dies lässt sich damit begründen, dass das Wissen über die Ursache fachdidaktisch betrachtet genauso wichtig ist wie das Fehlerbild selbst, weil sich eine wirkungsvolle Korrekturanweisung eben auch auf die Ursache des Fehlers beziehen sollte. Oftmals sind die Ursachen aber nicht zu beobachten, weswegen zunächst das beobachtbare Fehlerbild bekannt sein muss, ehe über mögliche Ursachen spekuliert werden kann. Auch die inhaltliche Qualität der Antworten fließt in die Bewertung mit ein. Wenn eine Beschreibung eine Zuordnung zu einem Fehlerbild erkennen lässt, dieses aber nicht exakt formuliert wurde, so wird die Antwort mit 1 BE bewertet.

Die inhaltliche Zuordnung der Fehlerbilder zu den geforderten Phasen Anlauf, Abwurf und Abfangen gestaltet sich schwierig, weil die Phasen fließend ineinander übergehen. Auch wenn in Frage 1 formuliert ist, dass die Wurfauslage nicht zum Anlauf (Frage 1), sondern zum Abwurf (Frage 2) gezählt wird, so sind doch immer auch Antworten möglich, die sich auf beide Phasen beziehen können. Beispielsweise ist das Setzen des falschen Stemmbeins nur das Ende einer langen Fehlerkette, die schon mehrere Anlaufschritte vorher zu fehlerhaften Bewegungsausprägungen führt. Daher wird festgelegt, dass Antworten, die sich auf den jeweiligen Bewegungsübergang zwischen den Phasen (das Setzen des Stemmbeins zwischen Anlauf und Abwurf sowie die Wurfarmbeschleunigung zwischen Abwurf und Abfangen) beziehen, als gültige Antworten akzeptiert werden. So werden beispielsweise die häufig genannten Antworten *falsches Stemmbein* (von 19 Testpersonen genannt) oder *kein Stemma* (von 14 Testpersonen genannt) als richtige Antwort auf Frage 1 (Fehlerbilder des Anlaufs) gewertet, obwohl die Aufforderung in der Fragestellung Antworten bezüglich der Wurfauslage explizit für Frage 1 ausschließt. Zusammenfassend ergeben sich Fehlerkataloge für die drei Einzelfragen (ohne Ursachenbeschreibungen), die in Tab. 45 des Appendix eingesehen werden können.

Auch die Auswertung von Teil D wird im Codebook genau vorgegeben, um die vielen Antwortmöglichkeiten in gleicher Weise zu bewerten. Die richtige Sortierung aller Bilder wird mit acht Bewertungseinheiten (BE) bepunktet (korrekte Reihenfolge: 4 2 7 1 6 3 8 5). Wenn in der Reihenfolge ein Bild an der falschen Position steht, die Reihenfolge der anderen sieben Bilder aber stimmt, wird wie folgt entschieden:

- 7 BE, wenn das falsche Bild nur um eine Position von ihrer richtigen Position abweicht
 - Beispiel: 4 2 7 1 **3** 6 8 5 (7 BE)
 - Beispiel: 4 **7** 2 1 6 3 8 5 (7 BE)

- 6 BE, wenn das falsche Bild um zwei Positionen von ihrer richtigen Position abweicht
 - o Beispiel: 4 **1** 2 7 6 3 8 5 (6 BE)
- 5 BE, wenn das falsche Bild um mindestens drei Positionen von ihrer richtigen Position abweicht
 - o Beispiel: 4 2 7 6 3 8 5 **1** (5 BE)

Wenn in der Reihenfolge zwei Bilder an der falschen Stelle stehen, wird wie folgt vorgegangen:

- 4 BE, wenn die Bilder vertauscht wurden und eine richtige Ziffer dazwischenliegt
 - o Beispiel: 4 **1** 7 **2** 6 3 8 5 (4 BE)
- 3 BE, wenn ein Bild davon mindestens drei Positionen von der richtigen Stelle entfernt ist
 - o Beispiel: **1** 4 2 **6** 7 3 8 5 (3 BE)

Wenn in der Reihenfolge drei Bilder an der falschen Stelle stehen, wird die Aufgabe mit 2 BE bewertet:

- o Beispiel: **2** 4 7 6 3 **5** 8 **1** (2 BE)

Eine Bewertungseinheit wird verliehen, wenn vier Bilder in falscher Position der Reihenfolge genannt werden, keine Bewertungseinheit hingegen, wenn mehr als vier Bilder falsch sind.

- o Beispiel: **1** 7 **2** **6** 3 **4** 5 8 (0 BE)

4.2.5.2 Reliabilität

Die Reliabilität als Gütekriterium für die Messgenauigkeit bzw. Zuverlässigkeit der Ergebnisse kann für Teil A und Teil B mit Hilfe des Cronbach- α -Koeffizienten überprüft werden. Bezüglich der ermittelten part-whole-korrigierten Trennschärfe r_{ix} (über die Werte der korrigierten Item-Skala-Korrelation) lässt sich für die ursprünglich 32 Items von Teil A feststellen, dass die einzelnen Fragen eine große Streuung aufweisen (von -,169 bis ,351). Dies mag vor allem der Tatsache geschuldet sein, dass Aufgaben mit einer besonders hohen Schwierigkeit stets eine geringere Trennschärfe aufweisen als Items mit einer mittleren Schwierigkeit (Lienert & Raatz, 1998, S. 58). Der Empfehlung von Lienert und Raatz (1998, S. 115) entsprechend wurden vier Fragen mit negativer Trennschärfe a posteriori eliminiert, weshalb in die weitere Berechnung des Gesamtscores nur noch 28 Items eingingen. Durch den Ausschluss dieser vier Fragen konnte der Wert für Kronbachs Alpha deutlich erhöht werden (Bühner (2011, S. 128) spricht von alpha-Maximierung) und liegt nach Auswertung der 59 Probandinnen und Probanden bei $r_{\alpha} = ,566$. Eine Eliminierung weiterer Items mit geringer Trennschärfe hätte zu keiner großen Reliabilitätsverbesserung mehr geführt, dafür hätte aber möglicherweise die Entfernung inhaltlich anspruchsvoller Fragen zu einer Verzerrung des Gesamtergebnisses geführt, daher wurde die Auswertung mit 28 Items durchgeführt.

Der ursprüngliche Teil B (Biomechanisches Wissen) mit 16 Items erreichte nur einen Koeffizienten von $\alpha = 0,212$. Bei einem genaueren Blick auf die Item-Skala-Korrelation zeigte sich, dass gleich sechs Items einen negativen Trennschärfewert (zwischen -0,004 und -0,212) aufwiesen. Drei weitere Werte mit einer Trennschärfe zwischen 0,017 und 0,058 wurden eliminiert, so dass sich der Wert auf $\alpha = 0,554$ verbesserte, was angesichts von nur acht Items eine durchaus zufriedenstellende interne Konsistenz vermuten lässt. Die acht eliminierten Fragen führten zu einer Verringerung des durchschnittlichen

Testwertes auf 5,76 (SD = 1,29), was darauf hindeutet, dass die eliminierten Items eine geringe Aufgabenschwierigkeit haben und von den meisten Testpersonen richtig beantwortet wurden. Somit darf begründet angenommen werden, dass trotz der Löschung der acht Items der inhaltliche Anspruch des Teils B erhalten bleibt. Die Itemtrennschärfe der acht ausgewählten Testfragen liegt zwischen 0,149 und 0,366.

Die ermittelten α -Werte für die Teile A und B liegen zwar unter der in der Literatur oft zu findenden Grenze von $\alpha = ,700$ für die Anerkennung als intern konsistente Skala (siehe Schmitt (1996)), jedoch ist dieser Schwellenwert nicht objektiv begründbar. Zum einen bleibt festzustellen, dass ein hoher α -Wert noch kein alleiniger Beleg für einen fachdidaktisch gut konstruierten Test, für eine homogene Skala oder für hohe Korrelationen auf Itemebene ist. Zum anderen stellt Schecker (2014, S. 5) hierzu fest, dass vor allem bei inhaltlich anspruchsvollen Tests ein hoher α -Koeffizient nur schwer zu erreichen ist, ohne dabei Abstriche bei der inhaltlichen Ausrichtung der Items in Kauf zu nehmen. Die vollständige Abbildung der zu ermittelnden Testgröße und somit das Erreichen einer hohen Validität ist jedenfalls wichtiger als das Bemühen, einen hohen α -Wert vorweisen zu können. Auch der korrelative Zusammenhang zwischen Einzelitems, die man im Vorfeld als sehr ähnlich ansieht, fällt oftmals sehr schwach aus, obwohl die Items einer insgesamt hoch reliablen Skala angehören. Dies lässt die Schlussfolgerung zu, dass man die Ausprägung fachlichen Verständnisses nicht mit isolierten Items begründen kann, sondern dafür eine Itemgruppe heranziehen muss, was in Teil A durch die 28 Items bzw. in Teil B durch die 8 Items zu erreichen versucht wurde.

Die eben geäußerten Limitationen des Kronbachs Alpha kommen beispielsweise auch bei schulischer Leistungsmessung oder dem in dieser Studie verwendeten Fachwissenstest zum Tragen. Schecker (2014, S. 6) stellt hierzu fest, dass man ein Testverfahren somit auch bei niedrigen Reliabilitätswerten verwenden darf, wenn man gut begründen kann, dass eine bestimmte Auswahl an Testaufgaben die Kerninhalte und Kompetenzen eines Untersuchungsgegenstandes umfassend beschreibt. Dies wurde für den Fachwissenstest dieser Studie in den vorhergehenden Kapiteln 4.2.1 und 4.2.2 ausreichend zu begründen versucht.

4.2.5.3 Validität

Da es für die Validierung eines Tests kein standardisiertes Routineverfahren gibt, muss die Validität möglichst plausibel begründet werden, denn sie ist noch vor der Objektivität und der Reliabilität das grundlegendste der drei Hauptgütekriterien (Schmiemann & Lücken, 2014, S. 107f.). Das Fachwissen über die Wurfbewegung ist dabei ein latentes Merkmal und theoretisches Konstrukt, welches mittels eines Testverfahrens möglichst vollständig und gültig ermittelt werden soll. Für die Validitätsprüfung ist die theoretische Klärung des Fachwissens daher unbedingte Voraussetzung, um zu überprüfen, inwieweit die späteren Testergebnisse die Interpretation stützen. Für den Fachwissenstest wird dies in Form einer intensiven Literaturanalyse mit der Konstruktion eines Strukturmodells (siehe Abb. 33) für die Bereiche räumlich-zeitliches Wissen (siehe Kapitel 4.2.1), biomechanisches Wissen (siehe Kapitel 4.2.2), Fehlerwissen (siehe Kapitel 4.2.3) und visuelle Bewegungsvorstellung (siehe Kapitel 4.2.4) durchgeführt.

Die Eruiierung der vier Facetten des Fachwissens wirkt sich positiv auf die Inhaltsvalidität aus, weil dadurch sichergestellt ist, dass die Items alle relevanten Inhaltsbereiche abdecken (Repräsentationschluss) und nur unter Rückgriff auf das Fachwissen beantwortet werden können. Das Fachwissen über die leichtathletische Wurfbewegung ist überschaubar, so dass die Anzahl der Items nicht zu groß wird,

um einen das gesamte Fachwissen repräsentierenden Test zu entwickeln. Entsprechend der Empfehlung von Murphy und Davidshofer (1998, S. 115) wird im Interesse der Inhaltsvalidität darauf geachtet, bei der Item-Entwicklung festzulegen, welche Inhaltsbereiche durch welche Items gemessen werden, wie in Kapitel 4.2.1.2 (für das räumlich-zeitliche Wissen) und Kapitel 4.2.2.2 (für das biomechanische Wissen) nachvollzogen werden kann.

Die inhaltliche Überprüfung der Einzelitems von Teil A und Teil B erfolgte durch eine Expertenbefragung. Die Leichtathletiktrainer Roland Fleischmann (Landestrainer in Ruhestand) und Christoph Lehner (B-Lizenz-Inhaber mit Landestrainererfahrung) wurden bei der Testentwicklung zu Rate gezogen, um die Eignung der Items einzuschätzen und zu beurteilen, inwieweit diese das Konstrukt valide abbilden. Beide Experten waren in der Lage, die Items der Teile A und B fehlerfrei zu beantworten. Ein „Konsens der Kundigen“, wie er von Lienert und Raatz (1998, S. 11) proklamiert wird, konnte somit für jedes Item hergestellt werden, so dass keine Eliminierung von Items vorgenommen werden musste. Die Experten waren sich einig, dass kein wichtiger Inhaltsbereich vergessen wurde. Zusätzlich sind den Experten kleinere sprachliche und inhaltliche Ungereimtheiten aufgefallen, die im gemeinsamen Gespräch verbessert wurden. Die Herangehensweise der offenen Fragestellung zur Erkennung des Fehlerwissens in Teil C wurde diskutiert und als sinnvoll erachtet, wenngleich die Experten auf die Schwierigkeit der Auswertung hinwiesen. Auch die Bilder von Teil D konnten beide Experten innerhalb kurzer Zeit in die korrekte Reihenfolge zu bringen. Abschließend bleibt festzuhalten, dass die Testaufgaben der einzelnen Teile von Personen mit hohem leichtathletischen Fachwissen und Erfahrungsreichtum richtig beantwortet werden können.

Als letzte Stufe der Testvalidierung wurde der Fachwissenstest als Mini-Pilotstudie mit drei Sportstudierenden durchgeführt, die der späteren Probandengruppe sehr ähnlich sind. Probandin A hatte leichtathletische Vorerfahrungen als Athletin und Trainerin über das Studium hinaus, Probanden B und C kamen mit Leichtathletik lediglich im Laufe ihres bisherigen Studiums in Kontakt. Um die Wahrnehmung der Items durch die Probanden näher zu erforschen, wurde die Methode des lauten Denkens (Thinking-Aloud Processes nach Ericsson und Simon (1993, S. 78–82)) angewendet. Die Probanden verbalisieren dabei ihre Gedanken beim Ausfüllen des Testbogens laut, so dass der Testleiter auf Verständnisprobleme oder auch auf Beantwortungsmechanismen schließen konnte, um zu entscheiden, ob die Frage nachvollziehbar mit leichtathletischem Fachwissen beantwortet wurde. Die bisherige Testkonstruktion wurde insofern bestätigt, als dass alle drei Probanden alle Testaufgaben ohne Rückfragen oder inhaltlichen Unklarheiten beantworten konnten. Dies bestätigt das Urteil der Experten, dass die Testfragen und Items eindeutig formuliert sind. Überraschend war allerdings, dass Probandin A aufgrund ihres höheren Fachwissens wesentlich mehr Schwierigkeiten hatte und mehr Zeit brauchte, die Items und Fragenteile zu beantworten, weil jedes Item gedanklich nachvollzogen und reflektiert werden konnte. Formulierungen wurden viel öfter schon beim Durchlesen von Items infrage gestellt oder mit eigenen Konzepten verglichen. Insgesamt führte aber dieses vertiefte Beantworten zu mehr richtigen Antworten bei Teil A und B, was sich zumindest bei diesen drei Testpersonen eindeutig auf ein höheres Maß an Fachkenntnissen über die leichtathletische Wurfbewegung zurückführen lässt. Teil C wurde von Probandin A wesentlich ausführlicher beantwortet, während die Probanden B und C lediglich oberflächliche Fehlerbilder nennen konnte. Teil D hingegen konnte von allen drei Probanden in gleicher Weise fehlerfrei beantwortet werden. Bezüglich der Validität des Tests kann aufgrund der Beobachtungen während der Testbearbeitung begründet behauptet werden, dass es sehr wahrscheinlich erscheint, dass mit einem höheren Fachwissen über die leichtathletische Wurfbewegung auch ein besseres Testergebnis erzielt werden kann.

Die Gültigkeit des Testverfahrens kann auch über die Kriteriumsvalidität versucht werden, indem die Annahme eines Zusammenhangs der Testergebnisse mit einem externen Kriterium überprüft wird. Als externes Kriterium kann die Examensnote in der mündlichen Fachtheorieprüfung gelten, die im Rahmen der Staatsprüfung Leichtathletik abgelegt werden muss. Es kann die Hypothese überprüft werden, dass Probanden, die eine bessere Note in der mündlichen Fachtheorie bekommen (Notenwerte 1 bis 6), auch mehr Punkte im spezifischen Fachwissenstest über die leichtathletische Wurfbewegung erzielen. Um dies herauszufinden, wurde die Korrelation zwischen der Examensnote und dem Ergebnis der verschiedenen Testteile ermittelt. Problematisch daran ist, dass die Inhalte des Fachwissenstests nur einen geringen Teil der mündlichen Examensprüfung ausmachen und somit kein direkter Zusammenhang zwischen beiden Ergebnissen besteht. Zugleich erfolgt die Erhebung im Fachwissenstest mittels schriftlicher Fragen (zum Teil im Multiple-Choice-Format), die Examensnote beruht aber auf einer mündlichen Prüfungsleistung. Außerdem liegt das externe Kriterium nur bei 19 Probandinnen und Probanden vor, wobei acht Testpersonen diese Prüfung schon vor dem Testtermin abgelegt hatten und elf Testpersonen erst danach. Es handelt sich bei den acht Personen daher um eine retrospektive Validität und bei den elf Personen um eine Vorhersagevalidität. Für den Zusammenhang zwischen der Examensnote und dem Ergebnis von Teil A (Räumlich-zeitliches Wissen) beträgt der Korrelationskoeffizient für die Gesamtgruppe (N = 19) $r = -0,514$ mit einer Signifikanz (2-seitig) von $p = 0,024$.³⁴ Das negative Vorzeichen ist dadurch begründet, dass ein hohes RZ-Wissen durch einen hohen Zahlenwert ausgedrückt wird, während bei der Examensnote eine sehr gute Leistung mit einem niedrigen Zahlenwert (bestenfalls der Note 1) bestätigt wird. Werte für den Korrelationskoeffizienten zwischen $0,3 < r < 0,6$ werden als mittel bezeichnet (Bortz & Döring, 2009, S. 606). Aufgrund der oben genannten Probleme darf dieser Wert demnach nicht überbewertet werden, kann aber als zusätzliches Argument gelten, welches die Validität der 28 Items von Teil A bestätigt. Für Teil B beträgt der Korrelationskoeffizient hingegen lediglich $r = -0,262$. Für den Gesamtscore der Teile A, B und D beträgt der Korrelationskoeffizient $r = -0,414$.

4.3 Test zur Kontrolle der Bewegungsvorstellung (TKBV)

Positive Effekte von Mentalem Training bei der Rehabilitation (z. B. nach Schlaganfällen, Kniegelenkstransplantationen oder Kreuzbandrissen) konnten mittlerweile wissenschaftlich nachgewiesen werden, indem Patienten Bewegungen ausschließlich mental durchführten, weil sie für ein aktives, motorisches Training zu immobilisiert waren. Es darf daher begründet angenommen werden, dass sich eine gut ausgeprägte Fähigkeit zur Bewegungsvorstellung auch auf die zielgerichtete Beobachtung von Bewegungen auswirkt. Die Messbarkeit der Bewegungsvorstellungsfähigkeit gestaltet sich jedoch schwierig, da es sich um ein schwer erfassbares Konstrukt handelt, für das zwar bereits verschiedene Messverfahren entwickelt wurden, die jedoch wegen verschiedenster Limitationen bisher keine allgemeine Anerkennung finden konnten. Der Controllability-of-Motor-Imagery-Test von Nishida et al. (1986), der von Naito (1994) weiterentwickelt wurde, beinhaltet erstmals den Versuch der Überprüfung, ob Probanden nach mental durchzuführenden Aufgaben überhaupt eine entsprechende Vorstellung entwickeln. Dazu musste die Versuchsperson eine Reihe von Instruktionen zu verschiedenen Körperteilbewegungen mental nacheinander ablaufen lassen, um zum Schluss dann die finale Körperpo-

³⁴ Betrachtet man die Probanden getrennt, so ergibt sich für die retrospektive Validität ein Korrelationskoeffizient von $r = -0,353$ (Signifikanz von $p = 0,391$), für die Vorhersagevalidität $r = -0,574$ (Signifikanz von $p = 0,065$).

sition zu reproduzieren. Über die Qualität der Endposition wurde dann auf die Fähigkeit zur Bewegungsvorstellung geschlossen. Schott (2013) entwickelte eine deutschsprachige Version dieses Tests (und nannte ihn Test zur Kontrolle der Bewegungsvorstellung, TKBV) und konnte diesen als valides und gut einsetzbares Instrument zur Erhebung der Bewegungsvorstellungsfähigkeit etablieren (siehe Kapitel 4.3.2). Daher soll der TKBV auch in dieser Studie zur Messung der allgemeinen Bewegungsvorstellungsfähigkeit zum Einsatz kommen, über welche die Fähigkeit zum Ist-Sollwert-Vergleich operationalisiert werden kann.

Die Überprüfung der Bewegungsvorstellungsfähigkeit wird als Fenster in kognitiv-motorische Prozesse angesehen [...]: Informationen müssen aus dem Langzeitgedächtnis abgerufen werden, im Arbeitsgedächtnis manipuliert werden, die tatsächliche Ausführung jedoch inhibiert werden. Dies erfordert zum einen ein gutes Arbeitsgedächtnis, zum anderen Bewegungserfahrung. (Schott, 2013, S. 671)

Um diese These zu überprüfen, konzipierte Schott (2013) den Test zur Kontrolle der Bewegungsvorstellung, welcher zwei große Teile umfasst: 1. Wiedererkennen (recognition) und 2. Freie Wiedergabe (free recall). Bei beiden Teilen muss der Proband fünf aufeinanderfolgende Instruktionen (Körperteilbewegungen) mental manipulieren, d. h. seine bestehende Vorstellung ergänzen, transformieren und wieder reproduzieren. Pro Instruktion wird die Bewegung eines Körperteils (Kopf, Arme, Beine, Rumpf) beschrieben, die nur mental nachvollzogen werden darf, ohne sich dabei aktiv zu bewegen. Nach der letzten Instruktion wird der Proband aufgefordert, so schnell wie möglich bei Teil 1 aus fünf verschiedenen Körperpositionen die richtige Endposition auszuwählen, bei Teil 2 die richtige Endposition selbst einzunehmen. Jeder Teil enthält zehn Aufgaben, wobei die Startbedingung immer gleich ist (der Körper steht aufrecht mit geschlossenen Beinen und nach unten hängenden Armen). Die fünf folgenden Bewegungsinstruktionen werden im Abstand von drei Sekunden gegeben und lauten für Aufgabe 1 von Teil 1 beispielsweise:

1. Machen Sie einen Schritt mit dem rechten Fuß um ca. 50 cm nach vorne.
2. Nehmen Sie Ihren linken Arm gerade nach oben.
3. Bringen Sie Ihren rechten Arm zur rechten Seite in die Waagrechte.
4. Drehen Sie Ihren Oberkörper um 90° zur rechten Seite.
5. Drehen Sie Ihren Kopf um 90° zur rechten Seite.

Die zur Auswahl stehenden Körperpositionen dieser Aufgabe von Teil 1 sind in Abb. 40 dargestellt. Die vollständigen Instruktionen des TKBV können im Appendix (Abb. 114 und Abb. 115) eingesehen werden.

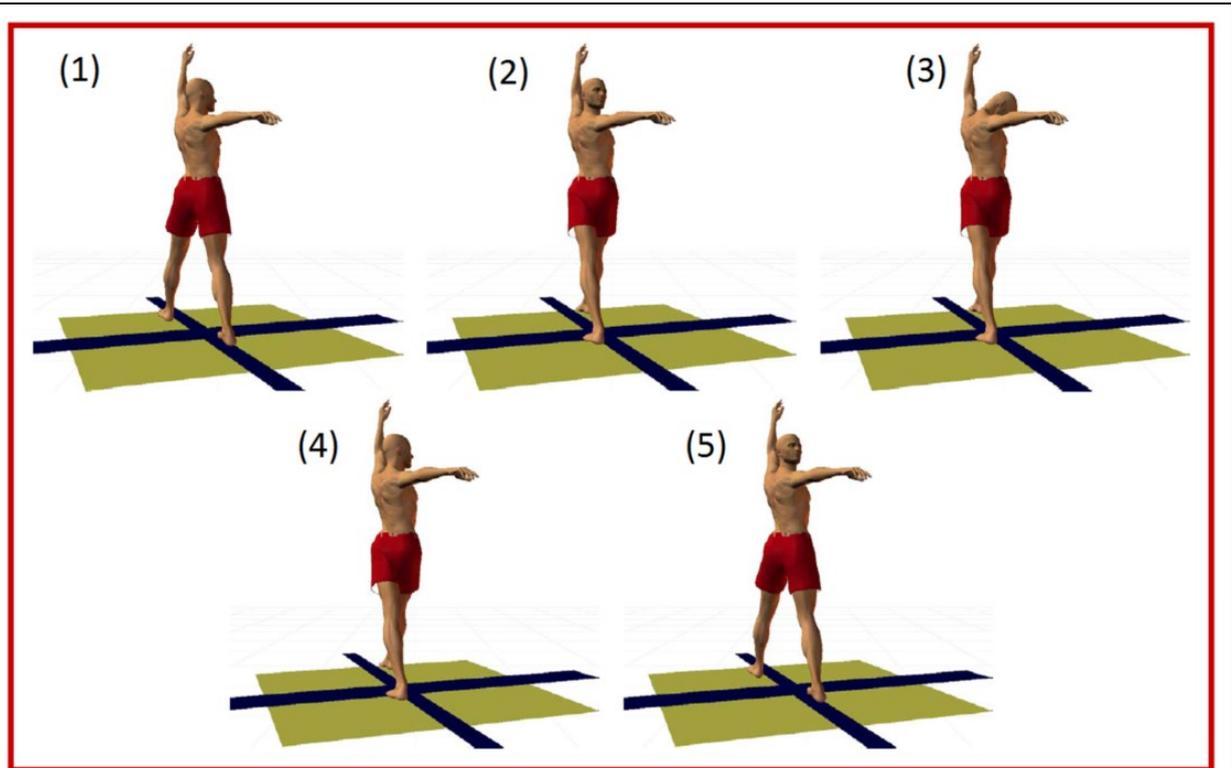


Abb. 40: Lösungsmöglichkeiten für Aufgabe 1 des TKBV-Teil 1 gemäß dem Original-Testmanual von Schott (die Nummerierung stellt eine Modifikation zur besseren Zuordnung der Lösungsmöglichkeiten im Rahmen dieser Studie dar)

4.3.1 Testdurchführung

Das Testmanual für den TKBV wurde von Frau Prof. Dr. Nadja Schott persönlich angefordert und die darin enthaltenen Vorgaben bezüglich ihrer Umsetzung diskutiert. Lediglich zum Zwecke der Testökonomie wurden kleinere Modifikationen vorgenommen, um eine größere Anzahl an Testpersonen gleichzeitig testen zu können. Teil 1 des TKBV (Aufgabe Wiedererkennen) wurde mit maximal 12 Teilnehmerinnen und Teilnehmern pro Testtermin (02.05., 10.05. und 17.05.2017 in jeweils zwei räumlich getrennten Gruppen) durchgeführt. Die fünf Instruktionen der jeweiligen Aufgaben wurden als Tonaufzeichnung wiedergegeben, die zur Auswahl stehenden Körperpositionen über Beamer auf einer Leinwand präsentiert. Für die Aufzeichnung der Testergebnisse wurde ein Single-Choice-Testbogen entwickelt (siehe Appendix Abb. 116, Aufgabe 1), in welchen die Testpersonen ihre favorisierte Lösung sowie die dafür benötigte Zeit eintrugen. Die Zeitmessung wurde mit Erscheinen der Körperpositionen gestartet und auch auf der Leinwand eingeblendet, so dass jeder Proband nach dem Ankreuzen seiner Lösung die benötigte Zeit in Sekunden ergänzen konnte. Außerdem wurden die Testpersonen darauf hingewiesen, dass bei Änderung der Lösung auch die Zeitangabe entsprechend korrigiert werden musste. Der standardisierte Test begann mit einer Probeaufgabe, welche von den Testpersonen mental gelöst werden musste, ohne jedoch schon ein Ergebnis in den Testbogen einzutragen. Nach einer kurzen Gelegenheit für Fragen und Unklarheiten konnten dann die zehn Aufgaben des ersten Teils des TKBV ohne weitere Rückfragen oder Unterbrechungen durchgeführt werden.

Teil 2 wurde dann von jeder Testperson an einem jeweils individuellen Termin absolviert, weil hierfür eine andere Testdurchführung notwendig war. Eine einzelne Testperson hörte wieder die digitalisierten Bewegungsinstruktionen an, bevor sie dazu aufgefordert wurde, möglichst schnell die entsprechende Körperposition selbst einzunehmen. Für die individuelle Auswertung musste pro Testperson ein Protokollant anwesend sein, der die Ergebnisse der Bewegungsausführung dokumentiert. Die Tests von Teil 2 wurden im Zeitraum vom 09.05. bis 24.05.2017 durchgeführt, wobei bis zu fünf Personen pro Termin gleichzeitig getestet werden konnten, indem fünf räumlich voneinander getrennte Testkabinen aufgebaut wurden, in denen die Testpersonen ohne Ablenkung durch andere die Aufgabe bewerkstelligen konnten. Bei richtig eingenommener Körperteilposition wurde das entsprechende Item auf dem Testbogen vom Protokollanten angekreuzt. Außerdem wurde von diesem gemäß TKBV-Testmanual die Zeit bis zum Einnehmen der endgültigen Körperposition gestoppt und festgehalten, ob die Endposition in mehreren Bewegungen nacheinander oder gleichzeitig (durch eine synchrone Bewegung) eingenommen wurde.

4.3.2 Gütekriterien

Die Güte der Testdurchführung des TKBV im Rahmen dieser Studie kann ähnlich wie beim Fachwissenstest mithilfe der drei Hauptgütekriterien beurteilt werden. Auf die Durchführungsobjektivität wurde speziell geachtet, indem das TKBV-Manual von Schott (2013) in möglichst vielen Punkten umgesetzt wurde. So wurden beispielsweise die allgemeinen und spezifischen Instruktionen aus dem Testmanual (siehe Abb. 117) vor Testbeginn wortwörtlich vorgelesen. Wie schon in Kapitel 4.3.1 erwähnt, wurden die Instruktionen der konkreten Aufgabenstellungen über eine Sprachaufnahme digitalisiert und bei der Testdurchführung abgespielt. Die Bilder der verschiedenen Körperpositionen wurden im Original übernommen und ebenfalls digitalisiert, um sie den Probanden sofort nach Beendigung der letzten Bewegungsanweisung der gesamten Testgruppe gleichzeitig zu präsentieren. Beide Maßnahmen sind Abweichungen zum Originaltest, bei dem den Testpersonen die Instruktionen vorgelesen und die Bilder in Papierform präsentiert wurden, wobei hier die Meinung vertreten wird, dass die Durchführungsobjektivität sich dadurch sogar noch erhöht. Die Auswertungsobjektivität wurde wiederum durch die Verwendung der automatisiert ablaufenden Auswertungssoftware KLAUS sichergestellt. Um Zweideutigkeiten bei der späteren Testauswertung zu vermeiden, erhielten die Probandinnen und Probanden eindeutige Anweisungen, wie sie bei der ein- oder mehrmaligen Änderung der Lösung vorzugehen haben. Bei Teil 2 des TKBV wurden die Protokollanten vor Testbeginn in ihre Beobachtungsaufgabe entsprechend eingewiesen und aufgefordert, bei Unklarheiten bezüglich der Auswertung den Testleiter um Hilfe zu bitten, um eine größtmögliche Auswertungsobjektivität zu erreichen. Gemäß Testmanual wurde in jeder Testkabine ein großes Kreuz auf dem Boden markiert, um dem Protokollanten die Beurteilung der entsprechenden Teilbewegungen (vor allem bezüglich der Bewegungsrichtung) zu erleichtern und somit Fehlerurteile zu vermeiden.

Die Rohdaten der einzelnen Testaufgaben wurden in einem zweiten Schritt in eine Excel-Tabelle übertragen, um eine Auswertung für jedes Item sowie für jeden Testteil separat durchführen zu können. Entsprechend der Auswertungsvorgaben des Testmanuals kann bei Teil 1 pro richtiger Bildauswahl ein Punkt und somit bei insgesamt zehn Testaufgaben ein Maximalscore von 10 Punkten erreicht werden. Bei Teil 2 gibt es pro Aufgabe fünf Bewegungsanweisungen, die richtig oder falsch eingenommen wer-

den können, und somit fünf Punkte pro Aufgabe. Bei insgesamt zehn Aufgaben ergibt sich ein Maximalscore von 50 Punkten. Hohe Punktwerte repräsentieren somit eine gute Bewegungsvorstellungsfähigkeit und deuten auf eine gut ausgeprägte Fähigkeit zum Ist-Sollwert-Vergleich hin.

Die theoretisch postulierte Struktur des TKBV konnte mittels Faktorenanalyse empirisch untermauert werden:

Mit den Faktoren Wiedererkennen (Erste-Person-Perspektive) sowie freie Wiedergabe (Dritt-Person-Perspektive) wurden knapp 42 % Varianz aufgeklärt. Beide Dimensionen weisen gute bis sehr gute psychometrische Eigenschaften auf, sowohl hinsichtlich der Reliabilität (z. B. interne Konsistenz, $\alpha = 0,89$ und $\alpha = 0,73$) als auch hinsichtlich der Ausschöpfung der Skalenbreite. (Schott, 2013, S. 670)

Zur Sicherung der Kriteriumsvalidität wurde der Zusammenhang zwischen der TKBV-Skala und dem Block-Tapping-Test (entwickelt von Schellig und Hättig (1993)), einem Testverfahren zur Erfassung der Leistungsfähigkeit des Arbeitsgedächtnisses, festgestellt, um zu überprüfen, ob ein von Schott (2012) festgestellter Zusammenhang zwischen der Leistung der Bewegungsvorstellungsfähigkeit und dem Arbeitsgedächtnis besteht. Gleicher Zusammenhang wurde auch zwischen der TKBV-Skala und den Probandenangaben zur eigenen körperlich-sportlichen Aktivität überprüft. Dabei konnte festgestellt werden, dass die TKBV-Skala hochsignifikant und deutlich sowohl mit der Leistung im Arbeitsgedächtnistest BTT (Teil 2 mit $r = 0,45$, Teil 1 mit $r = 0,38$) als auch mit der körperlich-sportlichen Aktivität (Teil 2 mit $r = 0,50$, Teil 1 mit $r = 0,36$) korreliert (Schott, 2013, S. 669). Somit ist auch für die Testpersonen dieser Studie, die allesamt Lehramtsstudierende mit Hauptfach Sport darstellen, ein hoher Gruppenschnitt bei den TKBV-Werten zu erwarten.

Ein Vergleich mit anderen Testverfahren zur Bewegungsvorstellung brachte nur eine bedingte inhaltliche Übereinstimmung zu Tage. Die Korrelation mit Testergebnissen des Movement Imagery Questionnaire (MIQ), welcher die kinästhetische und visuelle Vorstellungsfähigkeit ermittelt, liegt nur zwischen $r = 0,107$ für Teil 1 und $r = 0,143$ für Teil 2, die Testergebnisse der mentalen Chronometrie korrelieren etwas stärker ($r = -0,309$ für Teil 1, $r = -0,328$). Jedoch ist die Erhebungsmethode zu unterschiedlich, als dass hohe Korrelationswerte zu erwarten wären, wird doch bei der subjektiven Fragebogenerhebung des MIQ eher die psychische Leistungsfähigkeit als die tatsächliche Bewegungsvorstellungsfähigkeit bewertet. Schott (2013, S. 671) weist in diesem Zusammenhang darauf hin, „dass unterschiedliche Dimensionen mit der TKBV, der MIQ sowie dem TUG (Timed-Up-and-Go-Test³⁵, Anmerkung des Verfassers) geprüft werden [...], die darüber hinaus durch soziodemographische, psychosoziale und kognitive Komponenten maßgeblich beeinflusst werden können.“

³⁵ Der Timed-Up-and-Go-Test wurde für die klinische Anwendung konzipiert und misst die Zeit in Sekunden, die benötigt wird, um von einem Stuhl aufzustehen, eine Markierung in 3 m Entfernung zu umrunden, zurückzukommen und sich wieder hinzusetzen. Die Testpersonen müssen diese Aufgabe zunächst mental, dann praktisch vollziehen. Aus der Differenz der aktiven Zeit und der vorgestellten Zeit wird ein Index ermittelt, der die Qualität der Vorstellungsfähigkeit ausdrückt (Beauchet et al., 2010).

4.3.3 Exkurs: Test zur Bewegungsbeobachtung (BeBo) als Erweiterung des TKBV

Aufgrund der Ergebnisse von Schott (2013) ist zu erwarten, dass der TKBV für die Probandengruppe dieser Studien wenig Differenzierung hinsichtlich der Bewegungsvorstellungsfähigkeit liefert, weil jungen Personen mit einem hohen Umfang an körperlich-sportlicher Aktivität die Generierung und Manipulation motorischer Repräsentationen wesentlich besser gelingt als älteren Menschen mit wenig sportlich-körperlicher Aktivität. Daher wird ein den TKBV ergänzender Testteil entwickelt, der zur Überprüfung der Fähigkeit zum Bewegungsvergleich konzipiert ist. Im Unterschied zum TKBV wird nicht mit verbalen Instruktionen gearbeitet, sondern die Bewegungen werden als bewegtes Bild mittels Videoaufnahmen präsentiert. Zur Bewegungsdemonstration auf den Videos werden Personen mit einer hohen turnerischen Demonstrationsfähigkeit ausgewählt, weil durch eine hohe Qualität der ausgeführten Bewegungen der geforderte Bewegungsvergleich leichter gelingen kann. Pro Aufgabe muss sich die Testperson zwei Bewegungsvideos anschauen und in ihrer Bewegungsausführung vergleichen. In Video 1 führt ein Sportler (ähnlich wie beim TKBV) vier einfache Körperteilbewegungen (Beine, Arme, Kopf, Oberkörper) aus. Anschließend präsentiert der gleiche Sportler eine zweite Bewegung, welche von der Testperson mit der ersten Bewegungsausführung verglichen werden muss. Nach Betrachtung der Videos muss die Testperson mittels verschiedener Antwortmöglichkeiten auswählen, in welchen Bewegungsteilen sich Bewegung 2 von Bewegung 1 unterscheidet. Da das Vergleichen von zwei unterschiedlichen Bewegungsausführungen desselben Sportlers eine Beurteilungssituation darstellt, die im Rahmen einer Bewegungsrückmeldung bezüglich einer Verbesserung oder Verschlechterung der Bewegungsausführung im täglichen Trainings- und Unterrichtsalltag vorkommt (siehe Aufgabenteil C des Beurteilungstests in Kapitel 4.5.3), kann diese allgemeine Fähigkeit des Bewegungsvergleichs möglicherweise ein erster Gradmesser sein, ob dann auch der Bewegungsvergleich an einer spezielleren Bewegung wie dem leichtathletischen Schlagwurf gelingen kann.

Die Art der Aufgabenstellung des Bewegungsbeobachtungstests variiert jeweils in ihrem Komplexitätsgrad:

- Aufgaben 1 und 4: gleiche Reihenfolge mit einer falschen Teilbewegung
- Aufgaben 2 und 7: keine falschen Teilbewegungen, gleiche Reihenfolge
- Aufgaben 5 und 9: unterschiedliche Reihenfolge mit einer falschen Teilbewegung
- Aufgaben 3 und 8: unterschiedliche Reihenfolge mit zwei falschen Teilbewegungen
- Aufgaben 6 und 10: gleiche Reihenfolge mit zwei falschen Teilbewegungen

Außerdem werden vier unterschiedliche Kameraperspektiven verwendet, die zunächst innerhalb einer Aufgabenstellung gleichbleiben (z. B. A1, A2, A3 und A4), sich bei späteren Aufgaben aber auch zwischen den beiden Bewegungsdemonstrationen unterscheiden (z. B. A5, A6 und A7), so dass der Perspektivwechsel den Bewegungsvergleich zusätzlich erschwert (siehe Tab. 17). Die Testpersonen tragen ihre Antworten in einen Testbogen ein, der wiederum durch die Software KLAUS automatisiert ausgewertet wird.

Tab. 17: Kameraperspektiven der zehn Aufgaben des Bewegungsbeobachtungstests

	Bewegung 1	Bewegung 2
A1	Perspektive diagonal links vorne (dlv)	Perspektive diagonal links vorne (dlv)
A2	Perspektive frontal vorne (fv)	Perspektive frontal vorne (fv)
A3	Perspektive seitlich rechts (sr)	Perspektive seitlich rechts (sr)
A4	Perspektive seitlich rechts (sr)	Perspektive seitlich rechts (sr)
A5	Perspektive diagonal links vorne (dlv)	Perspektive diagonal rechts vorne (drv)
A6	Perspektive diagonal links vorne (dlv)	Perspektive diagonal rechts vorne (drv)
A7	Perspektive diagonal rechts vorne (drv)	Perspektive diagonal rechts vorne (drv)
A8	Perspektive diagonal links vorne (dlv)	Perspektive diagonal rechts vorne (drv)
A9	Perspektive diagonal links vorne (dlv)	Perspektive diagonal links vorne (dlv)
A10	Perspektive diagonal links vorne (dlv)	Perspektive diagonal links vorne (dlv)

4.4 Eigenrealisation

Meinel et al. (2015, S. 381) betonen, dass „das *eigene Bewegungskönnen* [...] eine Voraussetzung dafür [ist], dass ein entsprechendes ‚Bewegungsgefühl‘ für zeitlich-dynamische Parameter und einen ‚inneren Mitvollzug‘ einer Bewegung aufgebracht werden kann“, vor allem bei der Einschätzung von Bewegungsrhythmen. Somit kann begründet angenommen werden, dass die Fähigkeit zur eigenen Bewegungsausführung ein möglicher Einflussfaktor auf die Fähigkeit zur Beurteilung derselben sportlichen Bewegung sein kann. Die Eigenrealisation ist jedoch schwierig zu messen, weil es keine rein objektiven Merkmale gibt, um die Bewegungsqualität des Schlagwurfs zu beurteilen. Selbst biomechanisch messbare Daten wie Winkel, Bodenkontaktzeiten oder Geschwindigkeiten sagen für sich genommen noch nicht viel über die Bewegungsausführung aus. Hochleistungsathleten im Speerwurf unterscheiden sich beispielsweise mitunter selbst in den biomechanischen Messergebnissen so stark, dass allgemeingültige leistungsbestimmende Parameter der Wurfweite nur schwer zu bestimmen sind.

Trotzdem wird im Rahmen dieser Studie versucht, eine methodische Vorgehensweise zu entwickeln, welche die Qualität der Eigenrealisation auf möglichst vergleichbare Weise feststellt. Zu diesem Zweck wird die Eigenrealisation über die qualitative Einschätzung verschiedener Technikdetails sowie die Verknüpfung mit biomechanischen Messdaten, die aus einer 2-D-Videoanalyse gewonnen werden, operationalisiert. Das Geschlecht der Testpersonen bleibt dabei unberücksichtigt, da von der Annahme

ausgegangen wird, dass der Schlagwurf von Männern und Frauen in gleicher Weise qualitativ hochwertig ausgeführt werden kann und lediglich in einigen wenigen dynamischen Messparametern ein Vorteil der männlichen Probanden anzunehmen ist.

4.4.1 Qualitative Detailanalyse

Zur Analyse der Eigenrealisation werden die Bewegungsausführungen des Schlagwurfs aller 59 Testpersonen mit zwei Kameras aus der Seitperspektive gefilmt. Eine Kamera (Modell Panasonic) zeichnet in Echtzeit mit 50 HD-Bildern pro Sekunde auf, eine zweite Kamera (Modell Sony FS 700) im Hochgeschwindigkeitsmodus mit 200 HD-Bildern pro Sekunde. Aufgrund der besseren und stabileren Aufnahmebedingungen werden die Bewegungen in einer Turnhalle aufgezeichnet. Jede Testperson wird nach einer kurzen Erprobungsphase aufgefordert, mit einem Vortex-Wurfheuler zwei Schlagwürfe aus dem Fünf-Schritt-Rhythmus innerhalb einer vorgegebenen Anlaufstrecke auszuführen. Als Vorgabe wird lediglich verlangt, eine dynamische Wurfausführung zu zeigen. Für die weitere Auswertung werden die beiden Würfe eines Probanden separat analysiert und miteinander verglichen, um den besseren der beiden Versuche zu ermitteln und diesen für die Messung des Eigenrealisationsniveaus zu verwenden.

In einem ersten Analyseschritt wird die Wurfbewegung jeder Testperson vom Autor dieser Studie qualitativ beurteilt. Die Kriterien für diese Detailanalyse stützen sich auf einen Videoanalysebogen, welcher auf einem Lehrgang von Swiss Athletics (am 29.11.2011 in St. Gallen) für den Speerwurf erarbeitet wurde (Fäh & Straubhaar, 2011) sowie einem Technikanalysedokument mit zwölf Schlüsselpunkten der Speerwurfbewegung (Fuchser, 2013). Beide Dokumente gehen zurück auf das Model technique analysis sheet von Tidow (1996a, S. 60f.), welches insgesamt 55 Technikdetails mit einer dreistufigen Einschätzungsskala (+ o -) erfasst. Der Videoanalysebogen enthält immer noch 44 Beobachtungsmerkmale, welcher für diese Studie weiter auf 30 Beurteilungspunkte reduziert wurde. Durch die Tatsache, dass die Demonstration der Eigenrealisation durch die Testpersonen auf einen Fünf-Schritt-Rhythmus beschränkt wurde, entfallen die Beobachtungsdetails für die zyklische Anlaufphase. Da das Wurfgerät bei manchen Athletinnen oder Athleten bereits vor Wurfbeginn zurückgeführt war, wird auf eine Beurteilung der Wurfgerätrückführung verzichtet. Das Resultat und somit auch die Qualität der Rückführung ist aber in der Beurteilung der Wurfarmposition im Impulsschritt sowie später in der Wurfauslage erfasst, welche nur dann optimal eingenommen werden kann, wenn auch die Rückführungsaktion gelungen ist. Die Analyse beginnt demnach mit der Phase des Impulsschritts (über vier Beobachtungsmerkmale) und setzt sich über das Einnehmen der Wurfauslage (sechs Merkmale), die Körperposition und -aktion beim Setzen des Stemmschritts (sechs Merkmale), die Bogenspannung (sechs Merkmale), die Schlagposition (sieben Merkmale) bis hin zum Abfangen (ein Körpermerkmal) fort. In Tab. 50 des Appendix sind die 30 Beobachtungsdetails zur Operationalisierung der einzelnen Wurfphasen nochmals mit den Verbalbeschreibungen für die Qualitätsstufen 1 und 3 dargestellt.

Exemplarisch wird kurz erklärt, wie aus der optimalen Bewegungsausführung ab dem Einnehmen der Wurfauslage einzelne Bewegungsdetails operationalisiert werden, um zu einer qualitativen Bewegungsbeurteilung zu gelangen. Die nummerierten Bewegungsmerkmale beziehen sich auf Tab. 50 im Appendix. Das Einnehmen der Wurfauslage, beginnend mit dem Aufsetzen und der Belastung des Druckbeins (Zeitpunkt t_2), stellt dabei einen wichtigen Moment dar. Diese zentrale Beurteilungsposition sagt viel über die Dynamik und spätere Effizienz der Wurfbewegung aus. Der Wurfarm sollte dabei noch gestreckt (Bewegungsmerkmal (BM) 6) und unverändert auf Schulterhöhe (BM 5) weit hinter dem Körper sein, während das Stemmbein sich bereits vor dem Druckbein befindet (BM 10). Beim

Aufsetzen des Stemmbeins muss durch die aktive Druckbearbeitung das Eindrehen der unteren Extremitäten bereits deutlich erkennbar vor dem Oberkörper fortgeschritten sein, so dass der Rumpf noch eine möglichst große Aufdreh-Reserve besitzt (BM 14 wird erst in der Phase des Stemmschritts beurteilt). Der Schwungarm sollte sich immer noch lang vor dem Körper befinden, ehe er bei weiterhin gestrecktem Stemmbein (BM 16) gebeugt neben den Körper bewegt und fixiert wird, so dass er im Moment der Bogenspannung aus der Seitperspektive immer noch zu sehen ist (BM 18 wird erst in der Phase der Bogenspannung beurteilt). Durch einen explosiven Schwenk der Wurfarmseite in Wurfrichtung (BM 19) wird der Spannungsaufbau im Oberkörper eingeleitet und der Wurfarm mit hoher Ellenbogenführung (BM 24) beschleunigt, bis Brust- und Beckenachse frontal nach vorne zeigen (BM 20), während das Druckbein mit dem Schleifkontakt immer noch den Boden berührt (BM 22). Die Schlagposition kurz vor Verlassen des Wurfgeräts aus der Wurfhand zeichnet sich durch ein Abbremsen im Becken (Hüftknick) (BM 28), einen leicht nach vorne geneigten Rumpf (BM 29) sowie ein Nachschlagen des Wurfarms in Wurfrichtung (BM 27) aus, während die vorigen Merkmale von Stemm- und Druckbein immer noch unverändert zu erkennen sein sollten (BM 25 und BM 26).

Speziell die Beurteilung des Abfangens stellte eine Schwierigkeit dar, weil diese Phase sehr uneinheitlich ausgeführt wurde. Bei 26 Testpersonen war die Anlaufgeschwindigkeit so gering, dass überhaupt kein Abfangen über eine Bremsbewegung des Druckbeins nötig war, so dass entschieden wurde, die beiden zeitlich gesehen bewegungsbeendenden Merkmale ‚Druckbein aktiv bremsend‘ sowie ‚Rumpfform hinter dem Druckbein‘ eliminiert wurden. Anstelle dessen wurde das Abbremsen durch das gestreckte Stemmbein, welches den Körper kurzzeitig in den Ballenstand hebt oder die gesamte noch vorhandene Bewegungsenergie aufnimmt, für die Beurteilung des Abfangens herangezogen.

4.4.2 Biomechanische Datenerhebung

In einem zweiten Analyseschritt wurden mit Hilfe einer 2-D-Auswertung der Videoaufnahmen biomechanische Messwerte ermittelt, um zum einen die qualitative Einschätzung einiger der 30 Bewegungsdetails abzusichern (siehe Appendix Tab. 51, Spalte B.Mw). So kann beispielsweise die qualitative Einschätzung des Bewegungsmerkmals 5 (Wurfarm unverändert in Schulterhöhe) durch den biomechanischen Messwert #10 (Höhe des Wurfarms zum Zeitpunkt t1 in cm, Bezugspunkt Schultergelenk) oder des Bewegungsmerkmals 16 (Stemmbein gestreckt) durch den biomechanischen Messwert #18 (Stemmbeinwinkel zum Zeitpunkt t1 in Grad) überprüft werden. Zum anderen können aber auch neue Aussagen über die dynamische und räumlich-zeitliche Bewegungsqualität getroffen werden, wenn Messgrößen wie die Abfluggeschwindigkeit oder die Länge der Wurfgerätsbeschleunigung (pull distance) in die Analyse miteinfließen. Welche Messwerte sinnvoll erhoben werden können, wird zunächst mittels einer intensiven Literaturarbeit über die biomechanische Analyse der Speerwurfbewegung geklärt.

4.4.2.1 Biomechanische Analyse der Speerwurfbewegung

Schon Anfang der 1990er Jahre machte sich Harnes (1990a) Gedanken, wie biomechanische Messdaten in der Trainingspraxis von Speerwerferinnen gewinnbringend verwendet werden können, und betont zugleich, „daß [sic] wir uns mit den bisher gesammelten Zahlen noch in der Grundlagen-Forschungsphase befinden und so im wissenschaftlichen Sinne noch wenig ‚beweisen‘ können“ (S. 19).

Das Hauptproblem besteht darin, aus vielen möglichen Messvariablen die brauchbaren auszuwählen und die verschiedenen Zahlen im Sinne von gut oder schlecht zu bewerten. Harnes (1990a) entschied sich beispielsweise für die Abwurfgeschwindigkeit, den Abwurfwinkel, die Länge von Impulsschritt und Stemschritt, die Anlaufgeschwindigkeit sowie den Beschleunigungsweg. Um die aktuellen technischen Bewegungsausführungen des Speerwurfs von Hochleistungsathleten festzustellen, wurden regelmäßig die Würfe der Endkampfteilnehmer und -teilnehmerinnen zurückliegender Leichtathletikweltmeisterschaften biomechanisch analysiert. Auf die dabei verwendeten Messgrößen sowie deren Verwendung zur Unterstützung der Bewegungsbeurteilung wird im Folgenden eingegangen.

4.4.2.1.1 Bewegungstempo: Anlaufgeschwindigkeit

Um eine rein qualitative Einschätzung des Bewegungstempos zu überprüfen oder abzusichern, kann es hilfreich sein, die tatsächliche Anlaufgeschwindigkeit zu verschiedenen Anlaufzeitpunkten zu messen. Murakami et al. (2006, S. 68f.) kritisieren, dass bei vorherigen biomechanischen Analysen des Speerwurfs hauptsächlich die Abwurfparameter wie Abfluggeschwindigkeit oder Abflugwinkel gemessen wurden, die Anlaufgeschwindigkeit (approach velocity) hingegen nur eine untergeordnete Rolle spielte. Daher wurde in den biomechanischen Analysen der acht Weltmeisterschaftsfinalisten von 2005 sowie von 49 japanischen Speerwerfern die Geschwindigkeit des Körperschwerpunkts unmittelbar vor dem Aufsetzen des Druckbeins (hinteres Bein) zur Wurfauslage bestimmt. Dabei konnte eine signifikante positive Korrelation zwischen der Endgeschwindigkeit des Werfers und der Wurfweite nachgewiesen werden ($r = .742$, $p < .001$).

Murakami et al. (2006) bestätigen die Messungen von Harnes (1990b, S.20), wonach Weltklassewerferinnen ein höheres Anlauftempo realisieren als Nachwuchswerferinnen. Außerdem unterscheidet sich auch der Geschwindigkeitsverlauf bis zum Abwurf. Während die Weltklasseathletinnen es schaffen, die hohe Anlaufgeschwindigkeit im drittletzten Schritt durch den Impulsschritt weiter zu steigern und bis zur Wurfauslage nur einen sehr geringen Geschwindigkeitsverlust zuzulassen, kommt es bei Nachwuchsathletinnen schon im Impulsschritt zu einem leichten Abbremsen der Anlaufgeschwindigkeit, die sich bis zur Wurfauslage nochmals weiter deutlich verringert. Harnes (1990b, S.20) schließt daraus, dass die Größe des Geschwindigkeitsverlusts zwischen v_{t3} (Geschwindigkeit im Moment des drittletzten Bodenkontakts) und v_{t1} (Geschwindigkeit im Moment des Aufsetzens des Stemsbeins) in Prozent ausgedrückt „eine genaue Zahl für gute oder schlechte Werfer/Würfe darstellt“ (z. B. bei Wurfweiten über 70 m 2–5 % Geschwindigkeitsverlust, über 60 m 5–10 % und über 50 m 10–15 %). Um den Verlauf des Bewegungstempos in dieser Studie biomechanisch einzuschätzen, wurde daher auch die Anlaufgeschwindigkeit beim Aufsetzen des drittletzten Bodenkontakts (t_3) und beim Stemsbeinaufsatz (t_1) gemessen und miteinander verglichen. Aus den eben zitierten Ergebnissen lässt sich für trainierte Athletinnen und Athleten zwar schließen, dass es eigentlich nicht darum geht, die Geschwindigkeit in den letzten beiden Schritten zu steigern, sondern sie konstant oder die Verluste möglichst gering zu halten, etwas anders sieht es aber bei der Einschätzung des Geschwindigkeitsverlaufs der Testpersonen dieser Studie aus. Da der Anlauf nur aus fünf Schritten besteht und somit die Beschleunigung vor dem Impulsschritt relativ gering ausfällt, ist eine Temposteigerung bis zum Stemschritt durchaus möglich und wird auch von einigen Testpersonen realisiert (siehe Appendix Abb. 99). Diese Feststellung gilt es bei der Interpretation der ermittelten Messwerte bezüglich der Einschätzung des Bewegungstempos zu berücksichtigen (siehe Kapitel 5.2.3).

4.4.2.1.2 Bewegungsrhythmus: Impuls- und Stemmschritt

Die Einschätzung des Bewegungsrhythmus hängt maßgeblich von der räumlichen und zeitlichen Beurteilung des Impuls- und Stemmschritts ab, welche mit Zeit- und Längenmessungen zwischen den letzten drei Bodenkontakten biomechanisch sehr gut möglich ist. Harnes (1990b, S.19) hat schon 1990 speziell die Anlaufgestaltung anhand von Weltklassewerferinnen untersucht und folgende Tendenzen für Stemmschritt (L0), Impulsschritt (L1) und den dritt- und viertletzten Schritt (L2 und L3) festgestellt:

- „Die durchschnittliche Schrittlänge sollte ungefähr 95–98 % der Körperhöhe betragen
- L3 ist der zweitkürzeste Schritt von den vier letzten Schritten
- L2 ist kürzer als L1
- L1 ist der weiteste Schritt
- L0 ist der kürzeste Schritt
- Der Impulsschritt sollte mindestens 10 bis 15 % weiter als die durchschnittliche Schrittlänge sein [...]
- Der Stemmschritt sollte mindestens 10 bis 15 % kürzer als die durchschnittliche Schrittlänge sein“

Bradshaw, Burnett und Hume (2012, S. 42) untersuchten die zeitliche und räumliche Schrittgestaltung der WM-Finalistinnen von 2009, indem sie die drei Phasen Impulsschritt (CP = Crossover stride Phase), Stemmschritt (DP = Delivery Phase) und Abwurfphase (RP = Release Phase) unterscheiden (siehe Abb. 41). Der Impulsschritt beginnt beim Rechtswerfer mit dem linken Fußaufsatz (LC = Left foot of Cross step) und endet mit dem Fußaufsatz des rechten Beins (RC = Right foot Contact). Der Beginn des Stemmschritts schließt sich direkt daran an und endet mit dem linken Fußaufsatz zur Wurfauslage. Die Abwurfphase setzt sich dann fort bis zum Verlassen des Geräts von der Wurfhand. Dabei erreichten die Athletinnen beim Impulsschritt eine durchschnittliche Zeitdauer von 350 ms (SD = 66), beim Stemmschritt 198 ms (SD = 39) und beim Abwurf (RP) 138 ms (SD = 13). Analog dazu betrug die durchschnittliche Länge des Impulsschritts 1,88 m (SD = 0,31) und die des Stemmschritts 1,53 m (SD = 0,21). Harnes (1990a, S.22) berechnete bei seiner Anlauf-Analyse der beiden Spitzenathletinnen Felke und Renk das Verhältnis zwischen der Impulsschrittlänge und der Stemmschrittlänge und kam bei den Topwerferinnen auf einen Wert von 1,71 (Felke 206/120 = 1,71) bzw. 1,62 (Renk 212/131 = 1,62).

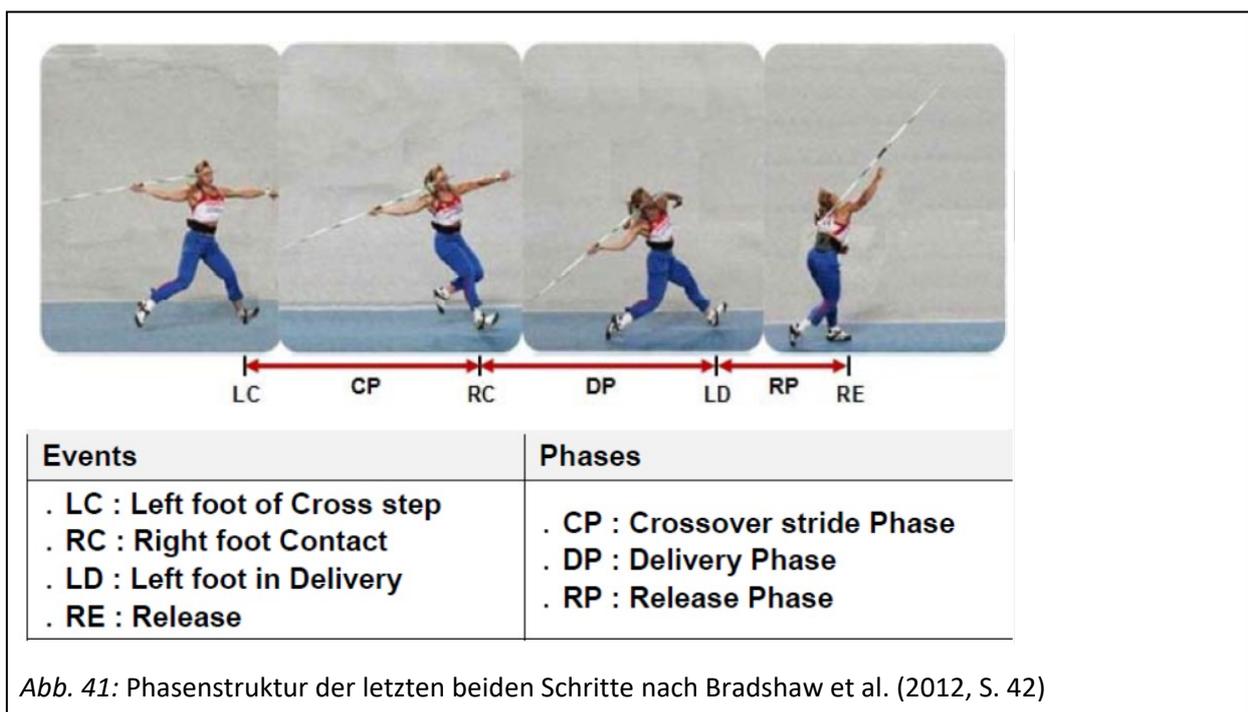


Abb. 41: Phasenstruktur der letzten beiden Schritte nach Bradshaw et al. (2012, S. 42)

In Anlehnung an Campos et al. (2004, S. 48) werden in dieser Studie vier Messzeitpunkte (siehe Abb. 42) festgelegt, die für die biomechanische Analyse der Abwurfbewegung interessant erscheinen und für eine zeitliche Vergleichbarkeit der Messdaten sorgen:

t_3 : Der linke Fuß (drittletzter Bodenkontakt) landet am Boden und markiert somit den Beginn des Impulsschritts

t_2 : Der rechte Fuß (Druckbein des Rechtswerfers) landet am Boden und markiert somit den Beginn des Stemmschritts (Preparatory Phase)

t_1 : Das linke Bein (Stemmbein des Rechtswerfers) landet am Boden und markiert somit den Beginn der Abwurfphase (Final Delivery Phase)

t_0 : Das Wurfgerät verlässt die Wurfhand

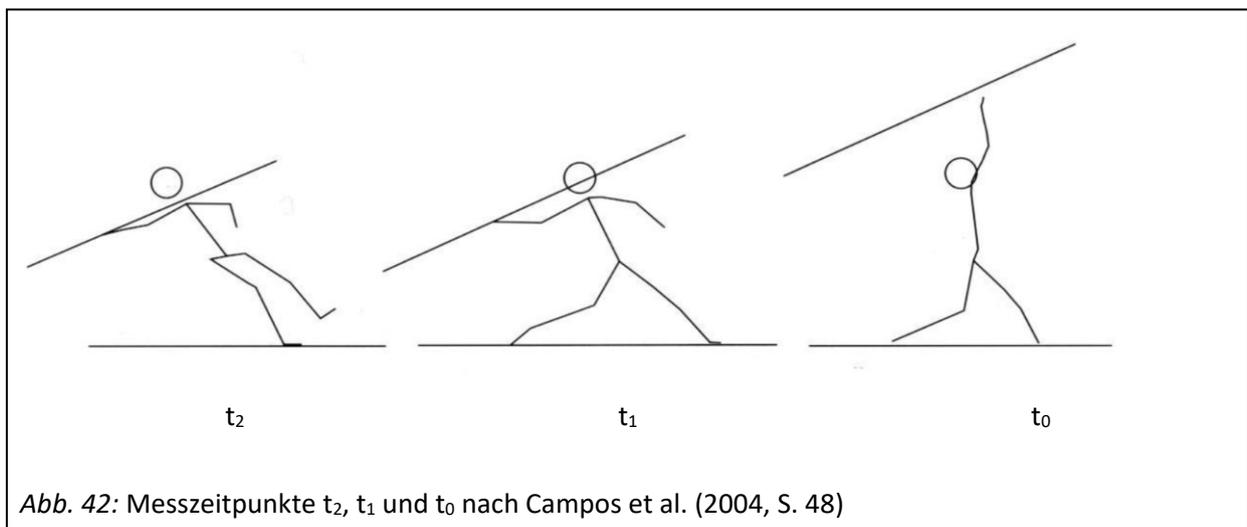
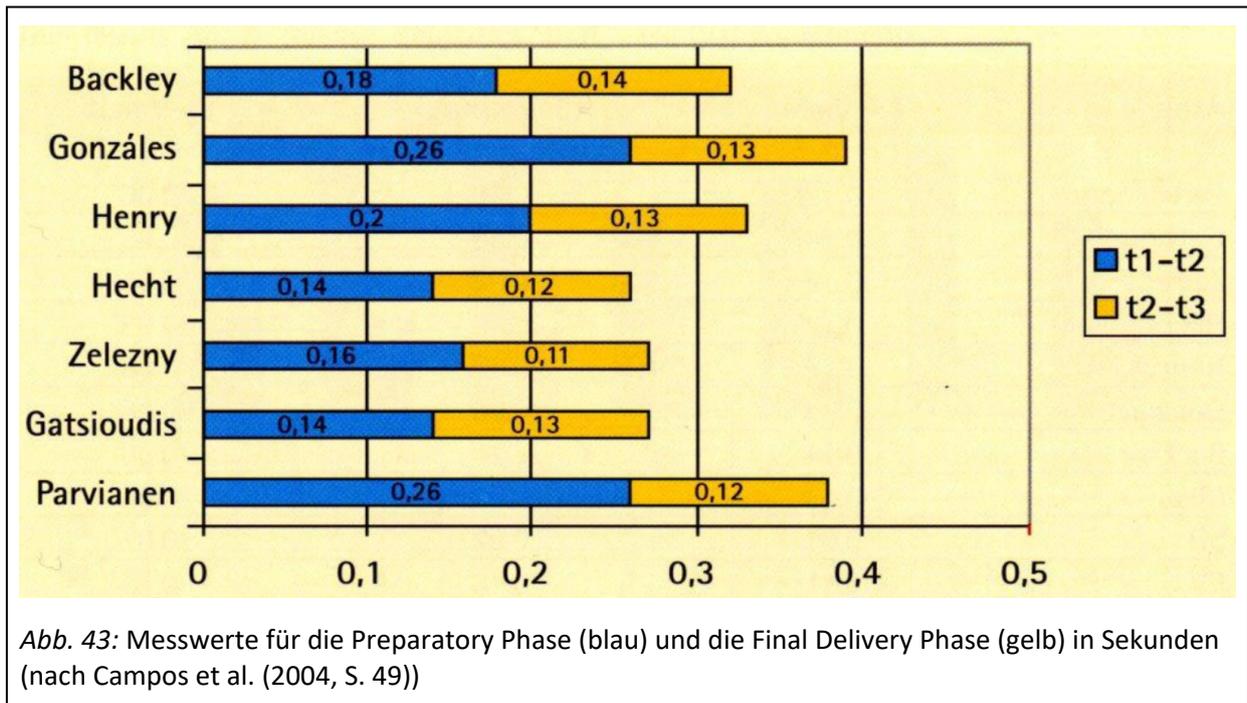


Abb. 42: Messzeitpunkte t_2 , t_1 und t_0 nach Campos et al. (2004, S. 48)

Mithilfe dieser drei Messzeitpunkte können folgende Zeitintervalle bestimmt werden:

- Die Dauer des Impulsschritts als Zeitintervall zwischen t_3 und t_2
- Die Dauer des Stemmschritts (Preparatory-Phase) als Zeitintervall zwischen t_2 und t_1
- Die Dauer der Abwurfbewegung (Final-Delivery-Phase, pull time) als Zeitintervall zwischen t_1 und t_0

Die Messungen bei den Finalteilnehmern des Speerwurfwettkampfs der Männer bei den Leichtathletik-Weltmeisterschaften 1999 ergaben, dass sich die größten Unterschiede zwischen den Topwerfern in der Preparatory-Phase ereignen. Die unterschiedlichen Athleten zeigten Intervalle von 140 bis 260 ms, während die Final-Delivery-Phase nur zwischen 100–140 ms dauerte. Es lassen sich also Werfer unterscheiden, die ihr Abwurftempo in einer sehr langen Preparatory-Phase (über 200 ms) verwirklichen (siehe Abb. 43: Gonzáles, Parvianen, Henry), während die andere Gruppe (Hecht, Zelezny, Gatsioudis) eine vergleichsweise kurze Preparatory-Phase (deutlich unter 200 ms) ausführte (Campos et al., 2004, S. 49).



4.4.2.1.3 Bewegungsumfang: Hüft- und Schulterrotation in der Wurfauslage und Wurfarmellbogen

Um den Bewegungsumfang biomechanisch abbilden zu können, muss der Beschleunigungsweg des Wurfgeräts operationalisiert werden. Dabei lohnt es sich, wieder einen Blick auf die Ausholbewegungen der Topwerfer zu richten, um ein besseres Verständnis für die biomechanischen Messdaten zu gewinnen. Um den Speer maximal zu beschleunigen, dreht der Werfer Hüfte und Schulter aus der senkrechten Position (90° = Schulter senkrecht zur Wurfrichtung) in die Wurfauslage (180° = Schulter parallel zur Wurfrichtung). Bei den Finalisten von 1999 fällt auf, dass die meisten Werfer zum Zeitpunkt t_2 die Hüfte auf Werte zwischen $135\text{--}151^\circ$ zurückgedreht haben und zum Zeitpunkt t_1 schon ein deutliches Nach-vorne-Drehen in der Hüfte realisiert haben. Das bedeutet also, dass die Hüftrotation nach vorne mit dem Setzen des Stemmbeins schon annähernd abgeschlossen ist. Die Schulterachse wird bei fast allen Athleten zum Zeitpunkt t_2 wesentlich weiter zurückgedreht, bei den meisten auf Werte zwischen $160\text{--}190^\circ$, ehe dann zum Zeitpunkt t_1 auch schon eine deutliche Beschleunigung nach vorne erfolgt, die aber der Hüfte um ca. $20\text{--}30^\circ$ hinterher läuft (Campos et al., 2004, S. 55, Table 6). Diese Messgrößen der Hüft- und Schulterrotation können zwar dazu dienen, den Bewegungsumfang der Ausholbewegung zu beschreiben, sind aber in dieser Studie messtechnisch aus der Seitperspektive der Videoaufnahmen nicht zu ermitteln.

Eine interessante und praxistaugliche Möglichkeit stellt die Messmethode von Campos et al. (2004, S. 55f.) dar, den Beschleunigungsweg (acceleration path) zu den beiden Messzeitpunkte t_2 (Preparatory Phase) und t_1 (Final Delivery Phase) wie folgt zu operationalisieren: Nach der Festlegung von Bartlett, Mueller, Lindinger, Brunner und Morris (1996) sowie Mero, Komi, Korjus, Navarro und Gregor (1994) wurde dabei der Abstand von der rechten Hüfte (des Rechtswerfers) bis zur vertikalen Projektion der Wurfhand gemessen (siehe Abb. 44). Auffällig war bei den Topwerfern, dass es zwischen t_2 und t_1 zu einer zwar geringfügigen, aber wahrnehmbaren Verringerung des Beschleunigungsweges kam. Zugleich bleibt festzuhalten, dass dieser Beschleunigungsweg neben der Rückrotation der Hüfte maßgeblich vom Beugewinkel des Wurfarms abhängt, weshalb dieser Messwert auch zur biomechanischen Beurteilung des Bewegungsumfanges herangezogen wird (siehe unten).

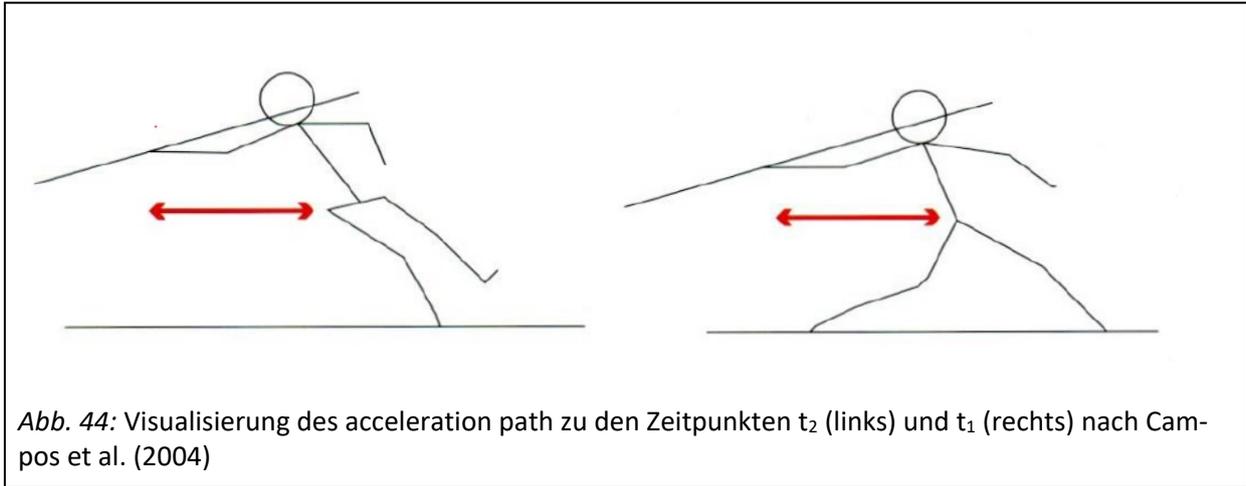


Abb. 44: Visualisierung des acceleration path zu den Zeitpunkten t_2 (links) und t_1 (rechts) nach Campos et al. (2004)

Murakami et al. (2006, S. 68) versuchen in ihren biomechanischen Analysen, den tatsächlichen Beschleunigungsweg des Wurfgeräts (die sog. pull distance) zu ermitteln, indem sie die Strecke der Wurfhand vom ersten Aufsetzen des Druckbeins in der Vorbereitung der Wurfauslage bis zum Verlassen des Speers aus der Wurfhand messen. Bei ihren Messungen konnte ein positiver und signifikanter Zusammenhang zwischen der pull distance und der Wurfweite hergestellt werden ($r = .426$, $p < .01$), während die dafür benötigte Zeitspanne (pull time) negativ mit der Wurfweite korrelierte ($r = -.418$, $p < .01$). Die WM-Finalisten von 2005 zeigten durchgängig eine größere pull distance und eine kürzere pull time als ihre Konkurrenten. Insofern vereint die pull distance zahlreiche auf die Maximierung des Beschleunigungsweges gerichtete Körperbewegungen in einem Messwert, der auch noch den Abwurfpunkt des Wurfgeräts miteinbezieht. Das Verhältnis zwischen pull distance und pull time kann außerdem einen ersten Anhaltspunkt für die Einschätzung der Wurfdynamik liefern.

Die Gelenkstellung des **Wurfarmellbogens** zu verschiedenen Zeitpunkten der Wurfbewegung ist häufig in biomechanischen Analysen zu finden. Aus technischer Hinsicht sollte der Wurfarm zum Zeitpunkt t_2 noch deutlich gestreckt sein und dann bei t_1 so stark als möglich gebeugt werden, um durch die Impulsübertragung maximal beschleunigt werden zu können. Fast alle Finalisten von 1999 erreichen Werte zwischen 158° und 172° bei t_2 , Beugewinkel zwischen 118° und 130° bei t_1 und dann wieder annähernd gestreckte Winkel zwischen 151° und 160° bei t_0 (siehe Abb. 45). Im Vergleich mit den weniger leistungsstarken japanischen Werfern zeigten die WM-Finalisten von 2005 einen verringerten Ellbogenwinkel im Moment des Abwurfs. Am effektivsten kann die Schulterrotation auf den Wurfarm übertragen werden, wenn der Ellbogen rechtwinklig gebeugt ist. Die Meinung vieler Trainer, dass der Ellbogen möglichst hoch gehalten werden und im Moment des Abwurfs möglichst stark gestreckt werden soll, muss demnach relativiert werden (Murakami et al., 2006, S. 69). Bradshaw et al. (2012, S. 42) bestätigen diese Ergebnisse: Die WM-Finalistinnen von 2011 haben einen durchschnittlichen Schulterwinkel relativ zur Waagrechten von $47,3^\circ$ ($SD = 10,1$) und einen Ellbogenwinkel relativ zur Waagrechten von $62,6^\circ$ ($SD = 10,6$).

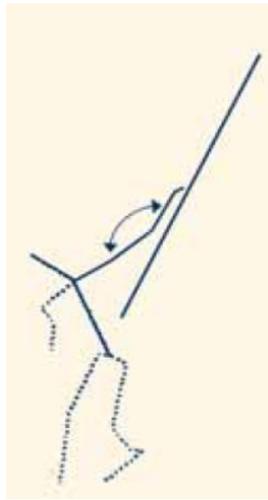


Abb. 45: Ellbogenwinkel beim Abwurf (Murakami et al., 2006, S. 77, Figure 10)

4.4.2.1.4 Bewegungskopplung: Beschleunigungsverlauf der Körperteilbewegungen

Die Bewegungskopplung stellt das biomechanisch am kompliziertesten zu erhebende Bewegungsmerkmal dar, weil es darum geht, die richtige Koordination von Teilimpulsen einzuschätzen. Campos et al. (2004, S. 50) halten diesbezüglich fest: „These sequential motions from the proximal to the distal segments are one of the fundamental keys to performance in over-arm throwing.“ Um die Qualität der Energieübertragung in der Final-Delivery-Phase biomechanisch zu überprüfen, wurden die Geschwindigkeitsverläufe von Hüfte, Schulter, Ellbogen und Wurfgerät gemessen. In Abb. 46 bestätigt sich der Theorieaspekt über das zeitliche Beschleunigen und Abbremsen genannter Gelenkpunkte. Beim Werfer Parvianen wurden Hüfte (grüne Linie) und Schulter (gelbe Linie) in gleicher Weise beschleunigt, die Hüfte jedoch stärker abgebrems als die Schulter. Der Ellbogen (pinke Linie) beschleunigt weiter und wird erst später abgebrems, so dass sich der Hauptimpuls auf den gehaltenen Speer (blaue Linie) überträgt. Die durchschnittlichen Messwerte der Finalwerfer von 1999 zeigen, dass das Maximum der Hüftbeschleunigung 130 ms, das Maximum der Schulterbeschleunigung 90 ms und das Maximum der Ellbogenbeschleunigung 60 ms vor dem Abwurfzeitpunkt des Wurfgeräts liegt. Auffällig dabei ist auch, dass bei allen Werfern außer einem das Maximum der Hüftbeschleunigung vor Zeitpunkt t_1 gemessen wurde, was darauf hindeutet, dass das Beschleunigen der Hüfte nach vorne noch vor dem Aufsetzen des Stemmbeins abgeschlossen ist (Campos et al., 2004, S. 50f.).

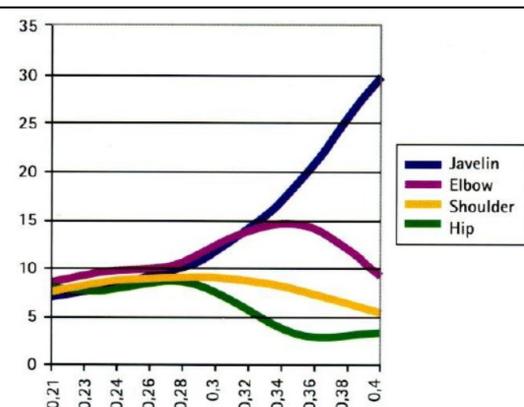


Abb. 46: Geschwindigkeitsverläufe beim Wurf von Parvianen (89.52) nach Campos et al. (2004, S. 50)

Da in dieser Studie die biomechanische Betrachtung lediglich nachträglich über eine zweidimensionale Videoanalyse erfolgt, kann die Qualität der Bewegungskopplung nicht durch konkret messbare Beschleunigungsverläufe abgesichert werden. Um jedoch indirekt zu beurteilen, ob die Bewegungskopplung dynamisch und räumlich-zeitlich effektiv ausgeführt wurde, kann der Parameter der Abfluggeschwindigkeit des Wurfgeräts herangezogen werden. Es herrscht Übereinstimmung in biomechanischen Untersuchungen, dass die Abfluggeschwindigkeit ein zentraler Parameter der Wurfleistung ist und von der Qualität der eben beschriebenen Abwurfkopplung abhängt.³⁶ Die Finalisten von 1999 erreichten dabei Werte zwischen 28,1 (für eine Weite von 85,43 m) und 29,7 m/s (für die Siegerweite von 89,52 m). Bei den WM-Finalistinnen von 2011 wurde eine durchschnittliche Abfluggeschwindigkeit von 25,60 m/s (SD = 1,16) gemessen, wobei die Siegerin (71,99 m) mit 25,11 m/s deutlich langsamer abwarf als die Zweitplatzierte (71,58 m) mit 26,17 m/s. Neben dem unterschiedlichen Abstand der Werferinnen von der Abwurflinie lässt sich eine geringere Abfluggeschwindigkeit demnach zumindest teilweise durch einen optimal ausgeführten Abwurfwinkel sowie eine maximale Abwurfhöhe kompensieren (Bradshaw et al., 2012, S. 43). Daher werden auch diese beiden Parameter in die biomechanische Analyse dieser Studie miteinbezogen.

Beim Abflugwinkel gilt im Speerwurf ein Bereich zwischen 32–37° als ideal, wobei die Finalisten von 1999 sehr unterschiedliche Werte zeigten (Hecht mit 40,1° und einer Wurfweite von 85,24 m, González mit 27,7° und 84,32 m und Parvianen mit 36,6° und der Siegerweite von 89,52 m) (Campos et al., 2004, S. 52). Auch bei der Weltmeisterschaft 2005 konnte nur ein nicht-signifikanter Zusammenhang mit der Wurfweite ermittelt werden (Murakami et al., 2006, S. 68). Messungen bei den Finalistinnen der Weltmeisterschaften 2011 ergaben einen durchschnittlichen Wurfwinkel von 38,0° (SD = 2,0°), wobei die Siegerinnen mit 39,4° eine Weite von 71,99 m erzielte, während die Viertplatzierte (65,24 m) lediglich einen Winkel von 33,2° realisierte. Zu große Abweichungen vom Idealbereich deuten aber auch auf eine schlechte räumlich-zeitliche Wurfkopplung hin.

Die Abwurfhöhe (siehe Abb. 47) stellt den dritten Faktor für die biomechanische Berechnung der Wurfweite dar, welche beeinflusst wird von der Körpergröße des Werfers, dem seitlichen Abknicken des Rumpfs und dem Kniewinkel des Stemmbeins. Grundsätzlich ist es biomechanisch vorteilhaft, wenn so hoch wie möglich abgeworfen wird bei gleichzeitigem Bodenkontakt des Stemmbeins. Hier zeigen aber selbst Spitzenwerfer unabhängig von der Körpergröße große individuelle Unterschiede (2,14 m bei Parvianen im Vergleich zu 1,80 m bei Zelezny). Murakami et al. (2006, S. 68) fanden lediglich eine nicht-signifikante Korrelation zwischen der Abwurfhöhe und der Wurfweite, weshalb dieser Parameter bei der biomechanischen Analyse nicht überbewertet werden darf.

³⁶ Campos et al. (2004, S. 51) konnte eine hohe ($r = 0,714$), aber nicht statistisch signifikante ($p = 0,072$) Korrelation zwischen der Wurfweite und der Abfluggeschwindigkeit des Speers nachweisen. Murakami et al. (2006, S. 68) konnte bei der Weltmeisterschaft 2005 ebenso eine positive Korrelation ($r = 0,889$, $p < .001$) nachweisen.

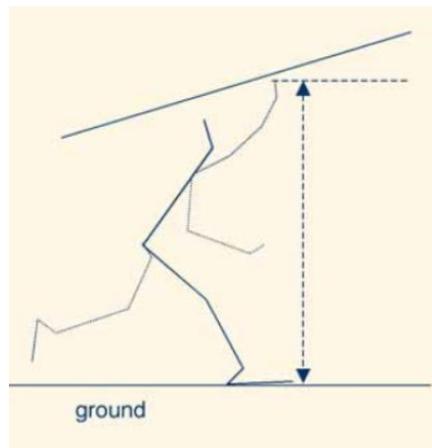


Abb. 47: Bestimmung der Abwurfhöhe nach Murakami et al. (2006, S. 74, Figure 5)

Ein letztes Detail, welches auf einen gelungenen Abschluss der Bewegungskopplung hindeutet, ist die Oberkörperposition zum Zeitpunkt des Abwurfs (t_0). Murakami et al. (2006, S. 69) betonen, dass es einen signifikanten Zusammenhang gibt zwischen der Wurfweite und der Vorwärtsbeugung des Oberkörpers beim Abwurf ($r = .463$, $p < .001$), der von der Senkrechten aus gemessen wird (siehe Abb. 48). Bradshaw et al. (2012, S. 42) konkretisiert dies mit konkreten Werten, wobei hier der Winkel von der Waagrechten aus gemessen wird (siehe Abb. 49): Der Oberkörperwinkel der WM-Finalistinnen von 2011 beträgt durchschnittlich $60,8^\circ$ ($SD = 8,3$), die Werte schwanken demnach sehr stark zwischen $73,5-46,7^\circ$ (die Siegerin realisierte einen Winkel von $61,2^\circ$). Es lässt sich daraus schlussfolgern, dass ein deutlich nach vorne geneigter Oberkörper auf eine gute Impulsübertragung der Anlaufgeschwindigkeit auf die Rumpfbeschleunigung hindeutet.

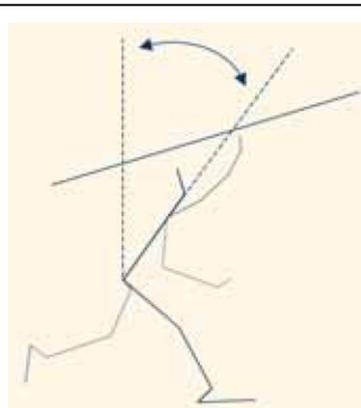


Abb. 48: Beugung des Oberkörpers beim Abwurf (Murakami et al., 2006, S. 79, Figure 14)

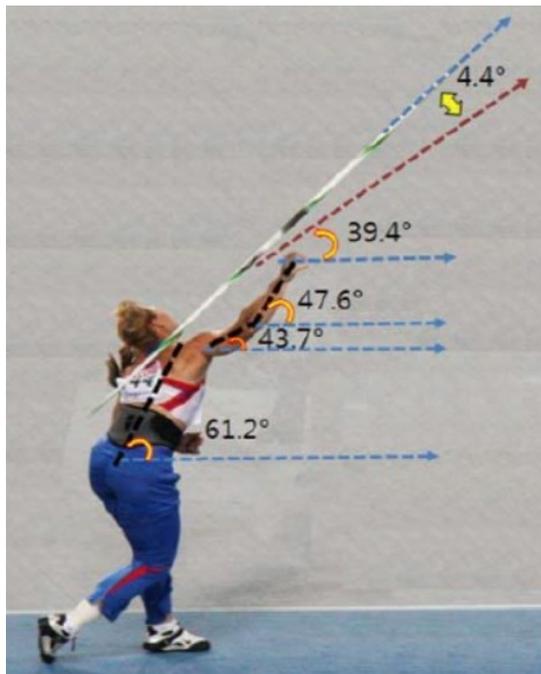


Abb. 49: Bestimmung von Oberkörperwinkel und Wurfarmwinkel nach Bradshaw et al. (2012, S. 43)

4.4.2.1.5 Bewegungspräzision: Kniewinkel des Stemmbeins in der Wurfauslage

Das Abbremsen der Wurfbewegung im Stemmschritt ist ein weiterer wichtiger Faktor, um eine maximale Abfluggeschwindigkeit zu realisieren und ein Übertreten zu vermeiden. Der Kniewinkel im Stemmbein (siehe Abb. 50) ist zwar ein selten erhobener Messwert bei biomechanischen Analysen, aber ein wichtiger Indikator für die Fähigkeit des Werfers, die horizontale Geschwindigkeit des Werferplus-Speer-Systems abzubremesen und auf das Wurfgerät zu übertragen, weshalb er in dieser Studie als Indikator für die Realisierung einer guten Bewegungspräzision dient. Die Wichtigkeit dieses Vorgangs mag besonders einleuchten, wenn man bedenkt, dass bei Spitzenathleten 60 Prozent der auf den Speer übertragenen Energie erst in den letzten 50 ms vor dem Abwurf erzeugt werden (Morriss & Barlett, 1996). Wenn ein Werfer es schafft, den Kniewinkel des Stemmbeins trotz Amortisation im Moment des Abwurfs des Wurfgeräts in einem Bereich zwischen 160–180° zu halten, so gilt dies als gelungene Stemmaktion. Besagte Finalisten von 1999 zeigen in diesem Bereich zwar sehr unterschiedliche Technikausführungen, dennoch lässt sich der Trend erkennen, dass das Knie bei t_1 fast gestreckt (zwischen 170–178°) aufgesetzt wird, dann zwischen t_1 und t_0 gebeugt wird (hier fallen die Unterschiede mit Werten zwischen 9–34° sehr stark aus), abschließend aber bis t_0 von den meisten Athleten wieder fast bis zum Wert von t_1 gestreckt wird (Campos et al., 2004, S. 54, Figure 6). Die Wichtigkeit des gestreckten Stemmbeins zeigten auch biomechanische Vergleichsuntersuchungen von Murakami et al. (2006) anhand der acht Weltmeisterschaftsfinalisten von 2005. Eine positive Korrelation konnte zwischen der Wurfweite und dem Kniewinkel des Stemmbeins nachgewiesen werden ($r = .310$, $p < .05$).

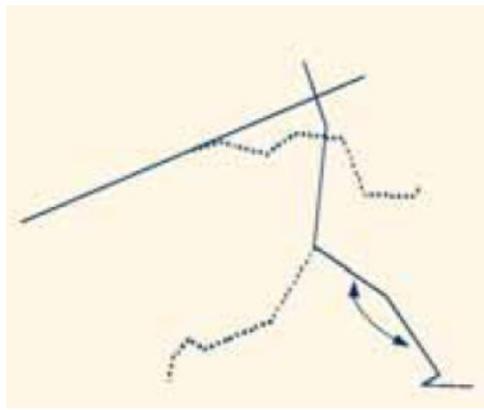


Abb. 50: Bestimmung des Stemmbeinwinkels nach Murakami et al. (2006, S. 78, Fig. 13)

4.4.2.2 Biomechanische Datenerhebung mittels Dartfish

Biomechanische Messungen erheben den Anspruch, auf physikalisch messbare Weise objektive Daten zu liefern, welche für die Interpretation der Bewegungsqualität hilfreich sein können. Da eine echte biomechanische Messung aufgrund des hohen Messaufwands in dieser Studie nicht möglich ist, wird auf die Auswertung von Highspeedaufnahmen mithilfe der Analysesoftware Dartfish zurückgegriffen, um eine Annäherung an objektiv messbare biomechanisch Daten zu erreichen. Die Tatsache, dass eine solche indirekte Messung mit manchen Ungenauigkeiten behaftet ist (wenn man allein schon die Linzenverzerrung bei der Kameraaufnahme bedenkt), führt aber nicht automatisch dazu, dass die erhobenen Werte unbrauchbar sind, sondern dass man sie entsprechend vorsichtig für mögliche Schlussfolgerungen einsetzen muss. Durch die Möglichkeit, die Bewegung in 200 Einzelbildern in HD-Qualität betrachten und analysieren zu können, kann jedoch eine sehr zufriedenstellende Güte der biomechanischen Datenerhebung erreicht werden. Dabei wird versucht, jede Messung mit einer möglichst hohen Standardisierung vorzunehmen, um die Vergleichbarkeit der Daten sicherzustellen. Zu diesem Zweck sind alle Bezugs- und Messpunkte in einem Testmanual beschrieben und festgelegt (siehe Appendix Abb. 118). Die wichtigsten Punkte werden im Folgenden für die Längen-, Winkel- und Zeitmessung dargestellt.

Um **Längenmessungen** durchzuführen, wird zuerst eine Kalibration des Längenmesstools mithilfe der Linierungen des Hallenbodens durchgeführt. Zur Eichung wird der äußere Abstand zwischen der hinteren Aufschlaglinie und der Grundlinie des Badmintonfeldes verwendet und genau in dem Bereich, in dem die Testperson anläuft, auf 76 cm eingestellt. Um Verkrümmungsungenauigkeiten im Kalibrierungsbereich zu vermeiden, wird die Kamera auf Höhe frontal zu diesen Linien aufgestellt und das Zentrum der Linse auf diesen Bereich scharfgestellt. Es bleibt zu erwähnen, dass bei Längenmessungen außerhalb des fovealen Bereichs der Kameralinse Verzerrungsungenauigkeiten auftreten, welche aber bei allen Testpersonen in gleicher Weise vorhanden sind, so dass innerhalb der Daten dieser Studie durchaus Vergleiche der Rohwerte möglich sind. Die Körpergröße der Testperson bleibt dabei unberücksichtigt, da sich bei der Datenanalyse gezeigt hat, dass vergleichsweise kleine Probandinnen und Probanden ihre fehlende Körpergröße mit einer technisch korrekten und dynamischen Ausführung der Wurfbewegung in allen Messwerten in einem für die Einschätzung der Bewegungsqualität ausreichenden Maß kompensieren können. Die Längenmessung ist ein zentraler Bestandteil der biomechanischen Erhebung, denn folgende Werte werden auf diese Weise bestimmt oder berechnet:

- Länge des vorletzten und letzten Anlaufschritts zwischen den jeweiligen Fußspitzen
- Abstand zwischen der Fußspitze des vorschwingenden Stemmbeins und der Fußspitze des bereits gesetzten Druckbeins zum Zeitpunkt t_2
- Abstand zwischen Hüfte und Wurfgerät (acceleration path) zu den Zeitpunkten t_2 und t_1
- Abstand zwischen Schulter und Wurfgerät zu den Zeitpunkten t_2 und t_1
- Abstand zwischen der Wurfgerätposition zu den Zeitpunkte t_2 und t_0 (pull distance t_2-t_0) sowie zu den Zeitpunkten t_1 und t_0 (pull distance t_1-t_0)
- Abwurfhöhe zum Zeitpunkt t_0
- Zurückgelegter Weg des Wurfgeräts nach t_0 (gemessen auf 10 Frames, entspricht einem Zeitintervall von 0,05 s) zur Berechnung der Abfluggeschwindigkeit des Wurfgeräts
- Zurückgelegter Weg eines festen Körperpunktes (Nasenspitze oder Ohrhöhle) zur Bestimmung der Anlaufgeschwindigkeit zu den Zeitpunkte t_3 , t_2 und t_1 (gemessen innerhalb des Zeitintervalls von Bodenberührung bis 10 Frames vorher)

Ein weiteres wichtiges Analysetool ist die **Winkelmessung**, welche nicht kalibriert werden muss. Jedoch ist auch hier wieder außerhalb des fovealen Bereichs eine Ungenauigkeit aufgrund der Linsenverzerrung zu akzeptieren. Zugleich befindet sich das zu messende Körpergelenk nicht immer frontal zur Kamera, wodurch ebenfalls Messungenauigkeiten auftreten. Bei der Messung des für die Beurteilung der Wurfbewegungen wichtigen Rumpfwinkels besteht außerdem die Schwierigkeit, dass die Gelenkpunkte nicht sichtbar sind und unter Beachtung der Körperstellung der Testperson geschätzt werden müssen. Ohne den Anspruch einer zu hohen Messobjektivität der Gelenkwinkel erheben zu wollen, können die Rohwerte aber trotzdem genutzt werden, um ein qualitatives Urteil von Teilbewegungen (z. B. Wurfarmstellung oder Stemmbeinwinkel) zu bekräftigen. In diesem Sinne werden die Winkelmessungen an folgenden Stellen erhoben:

- Oberkörperwinkel im Vergleich zur Vertikalen (Vor- oder Rücklage) zu den Zeitpunkten $t_{3\text{Ende}}$, $t_{2\text{Fußaufsatz}}$ und t_0
- Ellbogenwinkel zu den Zeitpunkten t_2 , t_1 und t_0 (bei t_0 speziell als Abweichung zur Vertikalen zur Beurteilung des Nachschlagens des Unterarms in Wurfrichtung)
- Winkel zwischen Horizontaler und Fußgelenk zum Zeitpunkt t_1 (zur Beurteilung des Fußaufsatzes über die Ferse)
- Abflugwinkel des Wurfgeräts (vom Bezugspunkt der Horizontalen, ermittelt vom Zeitpunkt t_0 bis 10 Frames später)

Die letzte Annäherung an biomechanische Messwerte erfolgt über die **Zeitmessung**. Es wird die Zeit zwischen den Bodenberührungen der letzten drei Bodenkontakte sowie der Zeit zwischen letztem Bodenkontakt und letztem Kontakt der Wurfhand mit dem Wurfgerät bestimmt. Die Messgenauigkeit des Zeitmesstools liegt bei 0,02 Sekunden, d. h. pro Frame erhöht sich die Zeitmessung um 0,02s. Da mit einer vierfachen Verlangsamung gefilmt wird, erfolgt eine entsprechende Umrechnung in Echtzeit, um die reale Dauer zwischen den Schritten sowie der Abwurfbewegung zu ermitteln. Folgende Messwerte fließen in die biomechanische Bewegungsbeurteilung mit ein:

- Zeitintervall zwischen dem drittletzten und vorletzten Bodenkontakt ($t_3 > t_2$) und dem vorletzten und letzten Bodenkontakt ($t_2 > t_1$)
- Zeitintervall zwischen letztem Bodenkontakt und letztem Kontakt der Wurfhand mit dem Wurfgerät ($t_1 > t_0$)
- Zeitintervall zwischen dem vorletzten Bodenkontakt und dem letzten Kontakt der Wurfhand mit dem Wurfgerät (pull time $t_2 > t_0$)

Neben der bereits angesprochenen Absicherung der qualitativen Beurteilung der Technikmerkmale werden sieben biomechanische Daten dazu verwendet, dynamische Bewegungsparameter (z. B. bezüglich der Anlauf- oder Abfluggeschwindigkeit) oder andere wichtige Messgrößen (z. B. den Abflugwinkel) in die Analyse miteinzubringen, da diese einiges über die Bewegungsqualität aussagen (siehe Kapitel 4.4.3) und es auch für die technische Beurteilung eine Rolle spielt, mit welcher Bewegungsdynamik die Wurfdemonstration ausgeführt werden kann. Die folgenden Nennungen der biomechanischen Messwerte mit Nummerierung beziehen sich dabei auf Tab. 51 im Appendix. Vor allem die biomechanischen Messwerte #1–3 (Anlaufgeschwindigkeiten zu den Zeitpunkte t_3 , t_2 und t_1) sowie die räumliche und zeitliche Ausprägung der letzten beiden Anlaufschritte (biomechanische Messwerte #4–5 sowie #6–7) zur Beurteilung des Anlaufrhythmus fließen in die Beurteilung der Eigenrealisation mit ein. Die Dynamik der eigentlichen Abwurfbewegung wird über die Parameter Pull distance #27 (t_2-t_0) und Pull time #29 (t_2-t_0) sowie die ermittelte Abfluggeschwindigkeit (biomechanischer Messwert #30) operationalisiert. Der gemessene Abflugwinkel des Wurfgeräts (biomechanischer Messwert #26) sowie die Abwurfhöhe in t_0 (biomechanischer Messwert #25) sind ebenfalls biomechanische Parameter, die nicht nur für die Ermittlung der theoretisch möglichen Wurfweite benutzt werden können, sondern auch Indikatoren dafür sind, ob die vielen Bewegungsdetails gut miteinander verknüpft wurden. Zugleich bleibt zu beachten, dass einige Parameter konstitutionsabhängig sind und somit nicht einfach in ihren Rohwerten miteinander verglichen werden können. So ist beispielsweise die Abwurfhöhe entscheidend abhängig von der Körpergröße, weshalb als Referenzwert die Differenz zwischen Abwurfhöhe und Körpergröße für die statistische Auswertung verwendet werden soll. In Tab. 51 des Appendix sind nochmals alle 30 biomechanisch erhobenen Messwerte aufgeführt. Die Zuordnung der biomechanischen Messwerte zur Absicherung der qualitativ beurteilten Bewegungsdetails kann genannter Tabelle in der Spalte B.Mw (Biomechanischer Messwert) nachvollzogen werden.

Technisch gesehen werden zunächst die Messzeitpunkte t_3 (drittletzter Bodenkontakt), t_2 (Druckbein setzt auf), t_1 (Stemmbein setzt auf) und t_0 (Wurfgerät verlässt die Wurfhand) bestimmt, wobei bei den Bodenkontakten jeweils die erste Berührung des Fußes (egal ob mit der Ferse oder der Fußspitze) als Beginn bzw. Ende der Zeitmessung definiert wird. Auf diese Weise können die Zeitintervalle für die letzten beiden Anlaufschritte (L2 = Impulsschritt, L1 = Stemmschritt) sowie die Zeit vom Aufsetzen des Druckbeins bis hin zum Abwurf des Wurfgeräts (sog. pull time von t_2 bis t_0) bestimmt werden. Zu beachten gilt, dass die ermittelte Zeit mit dem Faktor 0,25 multipliziert werden muss, um die vierfache Verlangsamung der Aufnahme auszugleichen. Ebenso kann die Schrittlänge von Impuls- und Stemmschritt über die Entfernungsmessung zwischen den jeweiligen Fußspitzen sowie die Beschleunigungsstrecke des Wurfgeräts (sog. Pull distance von t_2 bis t_0) bestimmt werden. Der Abflugwinkel des Wurfgeräts ist ebenso eine leicht zu erhebende Größe, weil die Flugbahn des Heulers nach Verlassen der Wurfhand auf dem Video verfolgt und der entsprechende Winkel relativ genau bestimmt werden kann. Etwas aufwändiger gestaltet sich die Messung der Abfluggeschwindigkeit, welche über die entsprechende Flugweite des Wurfgeräts vom Zeitpunkt des Verlassens der Wurfhand bis exakt zehn Frames später (entspricht dem Zeitintervall von $10 \times 0,005$ s) berechnet werden kann (siehe Abb. 51).

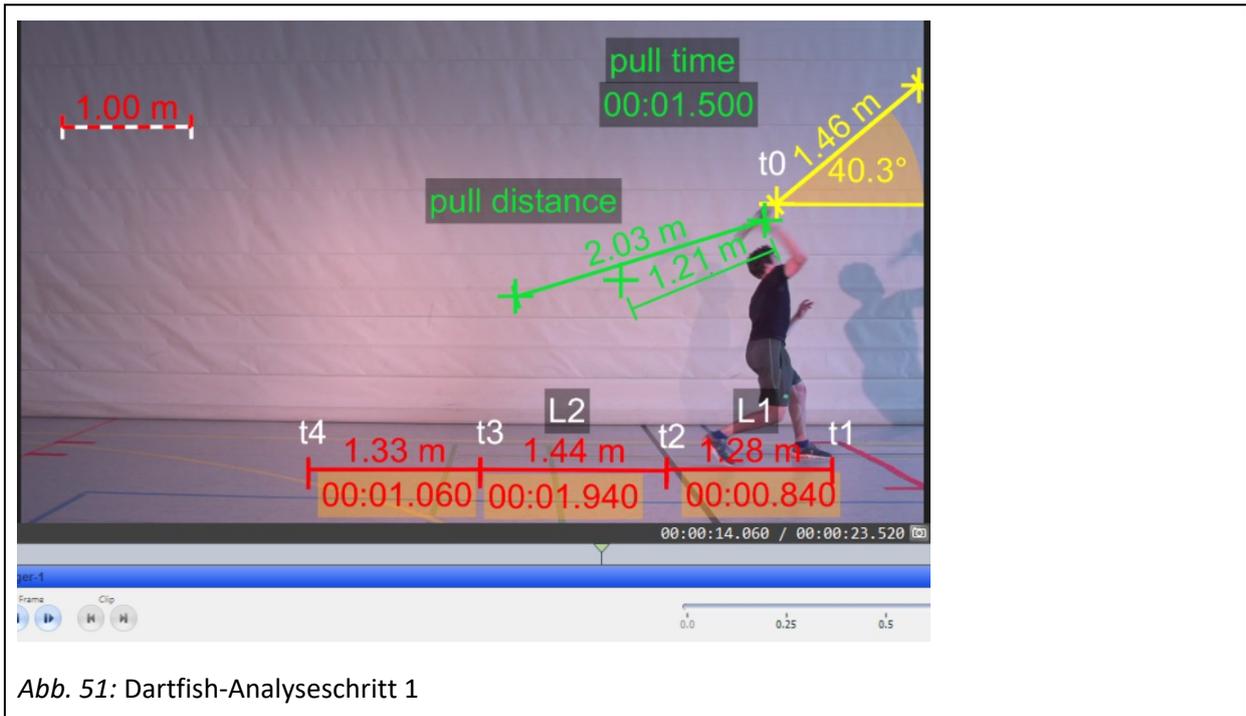


Abb. 51: Dartfish-Analyseschritt 1

In Analyseschritt 2 wird versucht, die Anlaufgeschwindigkeit so genau wie möglich zu ermitteln. Dazu wird ein ausgewählter Punkt nahe dem Körperzentrum (beispielsweise eine Markierung auf dem T-Shirt oder als Näherungswert auch das Ohrläppchen), der sich innerhalb des kurzen Messintervalls von zehn Frames (entspricht wieder dem Zeitintervall von 0,05 s) möglichst linear bewegt und nicht Teil einer sich gerade rotierenden Körperregion ist, verfolgt und über die zurückgelegte Strecke dieses Körperpunkts die momentane Geschwindigkeit berechnet. Die Messzeitpunkte für die Geschwindigkeitsanalyse sind wieder t_3 , t_2 und t_1 , wobei als Endpunkt der erste Bodenkontakt des jeweiligen Fußaufsatzes gewählt wird, von dem aus dann zehn Frames zurückgetrimmt wird, um die entsprechende Strecke des verfolgten Körperpunktes zu messen (siehe Abb. 52). Dadurch wird gewährleistet, dass die Geschwindigkeitsmessung in der Flugphase und somit noch vor dem Aufsetzen des möglicherweise bremsenden Fußes erfolgt.

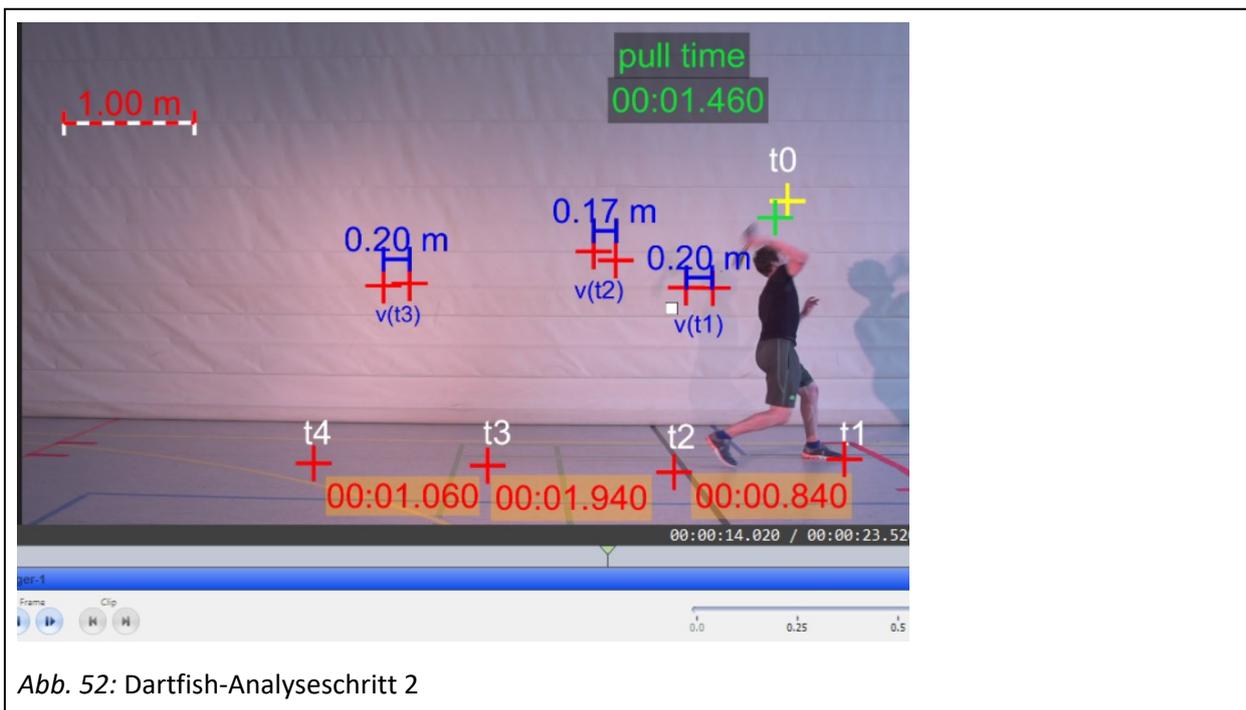


Abb. 52: Dartfish-Analyseschritt 2

In Analyseschritt 3 wird zunächst der Winkel des Wurfarmellbogens (vom Gelenkpunkt des Ellbogens an der Außenseite des Oberarms zur Achsel sowie an der Außenseite des Unterarms zum Handgelenk) zum Zeitpunkt des Aufsetzens des Stemmbeins gemessen (t_1). Je größer der Winkel ist, desto gestreckter ist der Wurfarm und desto länger der Beschleunigungsweg. Wenn der Ellbogenwinkel zum Zeitpunkt t_1 schon sehr klein ist, deutet dies auch auf eine schlechte Koordination der Teilimpulse hin, weil der Wurfarm zu früh schon nach vorne beschleunigt wurde. Außerdem wird der Winkel des Stemmbeins (vom Gelenkpunkt des Knies an der Außenseite des Oberschenkels zur Hüftbeuge sowie an der Außenseite des Unterschenkels zum Fußgelenk) zu den Zeitpunkten t_1 (Aufsetzen des Stemmbeins) und t_0 (Verlassen des Wurfgeräts von der Wurfhand) gemessen (siehe Abb. 53). Mit diesen beiden Werten kann überprüft werden, ob das Stemmbein gestreckt aufgesetzt wird und wie gut die Streckung während des Abwurfvorgangs aufrechterhalten und somit die Anlaufgeschwindigkeit abgebremst wird.

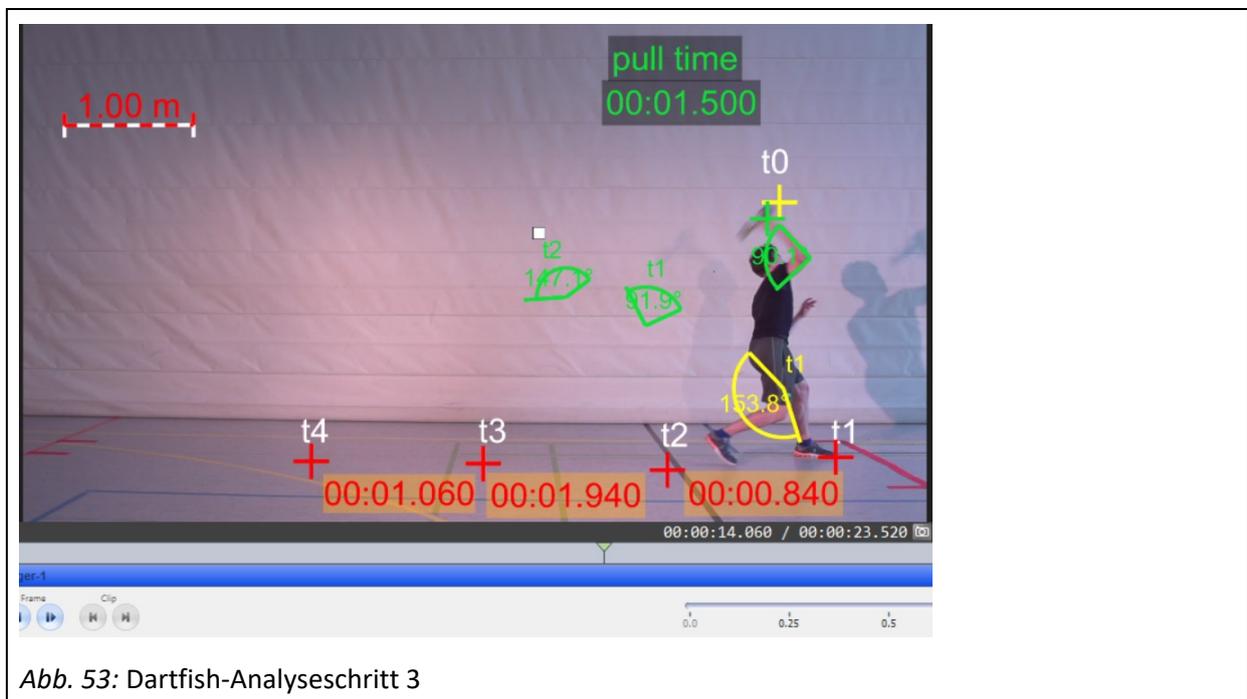


Abb. 53: Dartfish-Analyseschritt 3

In Analyseschritt 4 wird die Oberkörperposition zu den Zeitpunkten t_2 , t_1 und t_0 gemessen, wobei sich eine exakte Messung durch die Verwindung des Oberkörpers schwierig gestaltet. Der Winkel geht dann vom Gelenkpunkt am Hüftgelenk (Hüftbeuge) aus und wird zwischen den Schenkeln der Vertikalen und der Linie Hüfte-Schulter (Achsel) gemessen. Ein negativer Wert beschreibt eine Rücklage, ein positiver Wert eine Vorlage des Oberkörpers. Zudem wird die Abwurfhöhe als Entfernung zwischen Bodenmarker des Stemmbeins und Handposition beim Abwurf gemessen (siehe Abb. 54).

In Analyseschritt 5 wird der Wurfarm nochmals genauer betrachtet, indem die Entfernung zwischen der Wurfhand und der hinteren Hüfte der Wurfarmseite (acceleration path) zu den Zeitpunkten t_2 und t_1 gemessen wird. Diese beiden Messwerte deuten ähnlich wie der Ellbogenwinkel auf den Bewegungsumfang des Wurfarms sowie die vorhergehende Bewegungskopplung hin. Die Wurfarmhöhe zum Zeitpunkt t_2 (Aufsetzen des Druckbeins) als vertikale Entfernung zwischen der Wurfhand und der Schulterhöhe (Oberkante des Schultergelenks auf der Wurfarmseite) kann zur Beurteilung der hohen Wurfarmführung verwendet werden (siehe Abb. 55). Ein positiver Wert beschreibt, dass sich die Wurfhand

über Schulterhöhe befindet, ein negativer Wert deutet auf eine tiefe Wurfarmführung unterhalb des Schulterniveaus und somit einen hängenden Wurfarm hin.

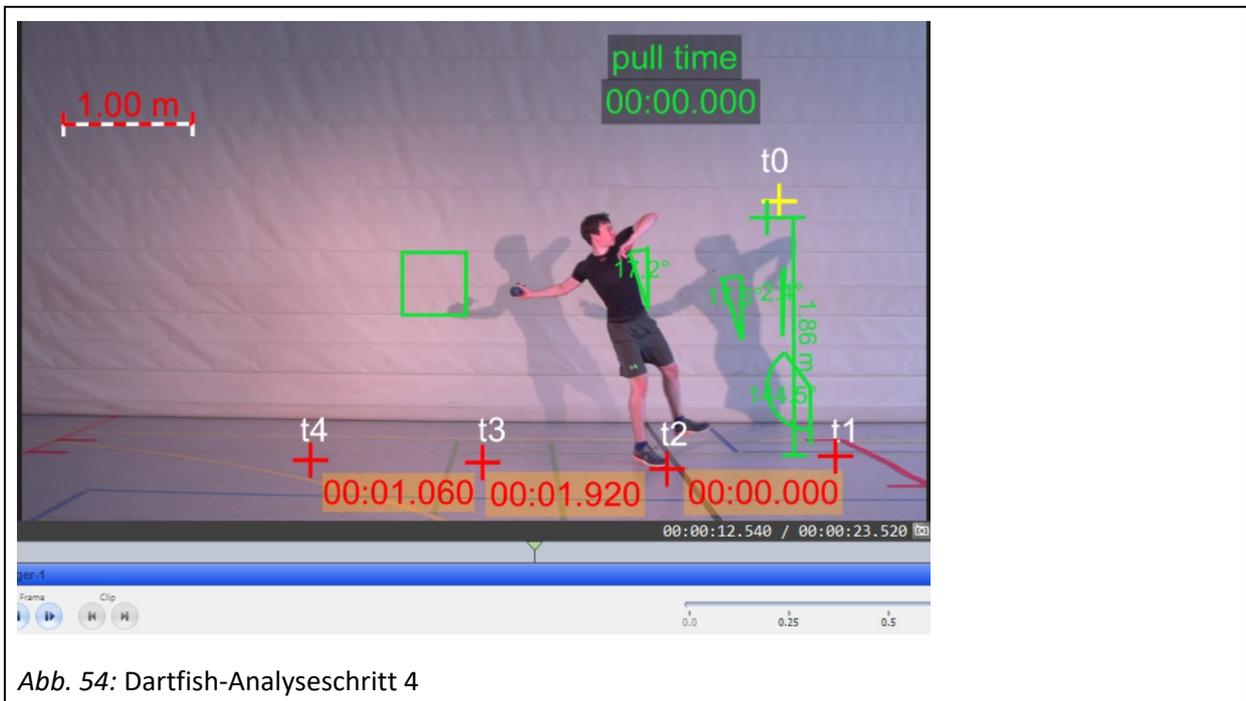


Abb. 54: Dartfish-Analyseschritt 4

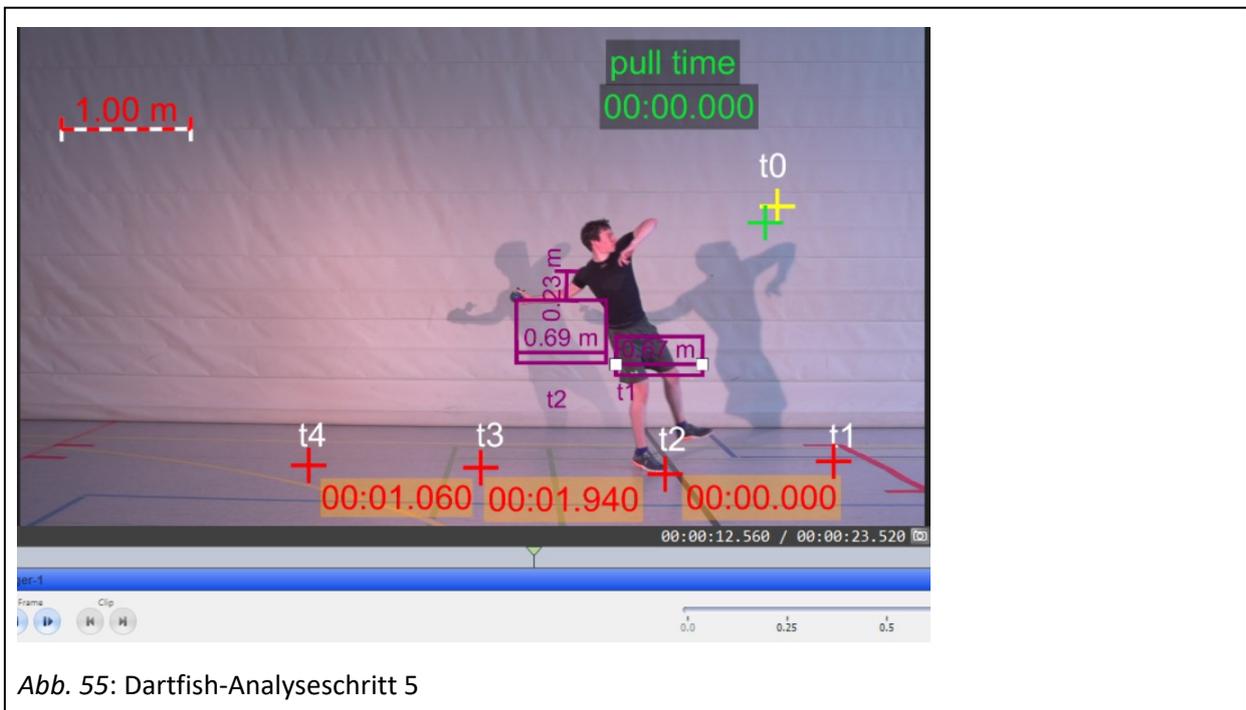


Abb. 55: Dartfish-Analyseschritt 5

Ein exemplarisches Datenblatt der 26 biomechanischen Messdaten einer Testperson, in welchem die Messwerte der Dartfish-Analyse eingetragen werden, kann in Tab. 52 des Appendix eingesehen werden.

4.4.3 Verwendung der biomechanischen Daten

Beide Erhebungsmethoden, die qualitative Beurteilung und die quantitative Messung, haben Vorteile und Nachteile. Daher wird auf nachvollziehbare Weise versucht, die beiden Datenbestände zusammenzubringen und daraus eine schlüssige Gesamtanalyse vorzunehmen. In einem ersten Schritt werden die biomechanischen Daten herangezogen, um die qualitative Einschätzung von einigen der 30 Beobachtungsdetails abzusichern (siehe Tab. 50 im Appendix, Spalte #BM).

Für die Phase des Impulsschrittes kann die Rückneigung des Rumpfes (Bewegungsmerkmal 2) über den biomechanischen Messwert #8 (Rumpfwinkel zum Zeitpunkt $t_{3\text{Ende}}$) und der flache Impuls des Abdruckbeins (Bewegungsmerkmal 4) über den biomechanischen Messwert #9 (Veränderung der Höhe des Körperschwerpunkts zwischen t_3 und t_2) kontrolliert werden. Die Wurfarmposition wird sowohl in der Wurfauslage (Zeitpunkt t_2 , Bewegungsmerkmale 5 und 6) als auch im Moment des Stemmbeinaufsatzes (Zeitpunkt t_1 , Bewegungsmerkmale 11 und 12) über die Messung der Wurfhandhöhe im Verhältnis zum Schultergelenk (biomechanischer Messwert #10 bzw. #15) und der Wurfarmstreckung im Verhältnis zur wurfarmseitigen Hüfte (biomechanischer Messwert #11 bzw. #16) sowie über die Ellbogenstreckung (biomechanischer Messwert #12 bzw. #17) beschrieben und beurteilt werden. Ebenso wird auch für den Zeitpunkt t_2 überprüft, ob der Rumpf eine entsprechende Rücklage aufweist (Bewegungsmerkmal 7, biomechanischer Messwert #13). Ob das Stemmbein sich dann bereits vor dem Druckbein befindet (Bewegungsmerkmal 10), wird über die Messung des Abstands der beiden Fußspitzen zueinander überprüft (biomechanischer Messwert #14). Ob das Stemmbein dann gestreckt (Bewegungsmerkmal 16) und über die Ferse (Bewegungsmerkmal 17) aufsetzt, wird über die biomechanischen Messwerte #19 (Winkel des Stemmbeins) und #20 (Winkel des Fußgelenks im Vergleich zum Boden) beschrieben. In der Schlagposition wird der Stemmbeinwinkel ebenfalls gemessen (biomechanischer Messwert #21) um zu überprüfen, ob das Stemmbein bis zum Abwurf gestreckt gehalten wird (Bewegungsmerkmal 25). Das Nachschlagen des Wurfarms in Wurfrichtung (Bewegungsmerkmal 27) wird mit der Messung des Winkels zwischen Unterarm und Vertikalen (biomechanischer Messwert #23) abgesichert, die Rumpfneigung im Moment des Abwurfs (Bewegungsmerkmal 29) über die Messung des Winkels zwischen Oberkörper und Vertikalen (biomechanischer Messwert #24).

In einem zweiten Schritt wird überlegt, welche biomechanischen Messwerte eine relativ objektive Einschätzung von zeitlichen-räumlichen und dynamischen Bewegungsverläufen ermöglichen, die auf qualitative Weise nur schwer zu beurteilen sind. Um das Bewegungstempo zu erfassen, werden die gemessenen Anlaufgeschwindigkeiten zu den Zeitpunkten t_3 und t_1 (biomechanische Messwerte #1 und #3) verwendet. Der Vergleich der beiden Messwerte lässt auch die Einschätzung zu, ob sich die Anlaufgeschwindigkeit zum Abwurf hin erhöht oder verringert. Das Verhältnis der Schrittlängen von Impuls- und Stemmschritt (biomechanische Messwerte #4 und #6) sowie deren zeitliches Verhältnis (biomechanische Messwerte #5 und #7) kann zur Deutung der Qualität des Anlaufrythmus herangezogen werden, denn auf diese Weise kann überprüft werden, ob der letzte Schritt wirklich kürzer ist als der vorletzte bzw. die letzten beiden Bodenkontakte schnell nacheinander gesetzt werden. Die zum Zeitpunkt t_0 erhobenen Parameter Abflugwinkel des Wurfgeräts (biomechanischer Messwert #26) und Abfluggeschwindigkeit (biomechanischer Messwert #27) lassen eine abschließende Bewertung zu, wie gut die Teilbewegungen des Abwurfs räumliche und dynamische miteinander gekoppelt wurden und die Anlaufgeschwindigkeit im Abwurf umgesetzt werden konnte. Da die Abwurfhöhe (biomechanischer Messwert #24) zwar für die exakte biomechanische Berechnung der Wurfweite nötig ist, sonst aber keine allzu große Auswirkung auf die Wurfweite hat (Murakami et al., 2006, S. 68), wird der erhobene Wert in dieser Studie nicht verwendet.

4.5 Beurteilungsfähigkeit der leichtathletischen Schlagwurfbewegung

Wie bereits aus den theoretischen Überlegungen über den Komplex der diagnostischen Kompetenz (siehe Kapitel 2.1.3) hervorgegangen, müssen für eine morphologische Bewegungsbeurteilung Kriterien festgelegt werden, die sinnesphysiologisch erfassbar sind und bewusstseinsfähig qualitativ eingeschätzt werden können. Dies scheint für die Meinel'schen Bewegungsmerkmale der Fall zu sein, weil in ihnen zahlreiche Bewegungsdetails in Überkriterien wie Bewegungsumfang oder Bewegungskopplung zusammengefasst werden, so dass eine Gesamteinschätzung des Bewegungsablaufs realistisch erscheint, ohne dass dazu eine unüberschaubare Anzahl an Einzelkriterien beobachtet werden müsste (siehe Kapitel 2.1.3.3.4). Im Rahmen der Konzeption eines Beurteilungstests gilt es nun, die ausgewählten sechs Beurteilungsmerkmale zu beschreiben (siehe Kapitel 4.5.1) und Überlegungen anzustellen, in welcher Form die Qualität ihrer Ausführung von den Testpersonen festgestellt werden kann (siehe Kapitel 4.5.2).

4.5.1 Beschreibung der Beurteilungsmerkmale

Obwohl eine wissenschaftliche Annäherung an den Komplex der Bewegungsbeurteilung nur über Einzeluntersuchungen möglich ist, besteht trotzdem weitgehend Einigkeit darüber, dass die Meinel'schen Bewegungsmerkmale als Ausdruck der Bewegungskoordination (Meinel, 1960) für die Bewegungsbeurteilung immer noch von zentraler Bedeutung sind. Die verschiedenen morphologischen Merkmale nach Meinel wurden in der Vergangenheit immer wieder als sogenannte „Beobachtungseinheiten“ (Neumaier, 1988) oder auch als „Kategorien zur Bewegungsbeobachtung und -beurteilung“ (Hartmann, 1999) herangezogen, um die zeitliche und dynamische Dimension einer Bewegung zu beschreiben, welche dem Beobachtenden kaum visuell zugänglich sind. So kann beispielsweise die zeitliche Komponente über die Geschwindigkeitsänderung (Bewegungstempo) oder die dynamische Komponente über die erzielte Wirkung von Krafteinsätzen (Bewegungsumfang oder Bewegungsstärke) relativ zuverlässig eingeschätzt werden. Die psychischen Aspekte, die über Ausdrucksformen der Verhaltensregulation oder sogenannte Handlungsmerkmale Aussagen zur Innensicht des Sichbewegenden gestatten und ebenfalls den Grad der erreichten Bewegungskoordination widerspiegeln können, werden in dieser Arbeit nicht behandelt, da sie für die originäre Bewegungsbeurteilung keine Rolle spielen. Stattdessen werden die sechs Meinel'schen Termini Bewegungsharmonie, Bewegungstempo, Bewegungsrhythmus, Bewegungsumfang, Bewegungskopplung und Bewegungspräzision als Beurteilungsmerkmale gewählt, um die Wurfbewegung in ihren vielen technischen Teilaspekten beobachtbar zu machen.

Die **Bewegungsharmonie der Gesamtausführung** wird hier als übergeordnetes Qualitätsmerkmal der Gesamtbewegung verstanden, ein „Kriterium für die Vollkommenheit der technischen Ausführung, für den Grad der Beherrschung der Technik, für den erreichten Stand in einem Lernprozess“ (Meinel et al., 2015, S. 122). Wenn eine Bewegung flüssig und koordiniert abläuft, sind die einzelnen Kraftimpulse optimal aufeinander abgestimmt. „Eckige Richtungsänderungen, Unterbrechungen oder Verzögerungen oder ein übermäßig harter, ruckhafter Krafteinsatz“ (Meinel et al., 2015, S. 122) hingegen sind Anzeichen für eine ungenügende Koordiniertheit der Muskelkraftimpulse. Im folgenden Beurteilungstest soll daher zu allererst eine Einschätzung über die grundsätzliche Koordiniertheit der gezeigten Bewegung getroffen werden, indem mittels dreier Qualitätsstufen die Bewegung in ihrer Gesamtausführung zwischen den Polen ‚flüssig, harmonisch und koordiniert‘ und ‚eckig, unrund und unkoordiniert‘

verortet wird. Dabei können durchaus auch konkrete Fehlerbilder einer Qualitätsstufe zugeordnet werden wie das Werfen über das falsche Stemmbein, welches bei Anfängern häufig auftritt und der Qualitätsstufe 3 (starke Abweichung hinsichtlich der Bewegungsharmonie) zugeordnet wird.

Die folgenden fünf Beobachtungskriterien sind dann an der chronologischen Bewegungsabfolge des Ballwurfs orientiert. Dabei wird versucht, jede wichtige Bewegungsphase mit einem relevanten Meinel'schen Bewegungsmerkmal zu verknüpfen. Die Einschätzung des **Bewegungstempos im Anlauf** bis hin zur Wurfauslage ist für das Resultat der Bewegung ein wichtiges Kriterium, weil die Vorbeschleunigung des Körpers und somit des Wurfgeräts durch den Anlauf einen entscheidenden Anteil an der Wurfweite hat. Dies kann sowohl die Geschwindigkeitsrelationen ganzer Bewegungsakte und ihrer Kombinationen sowie wesentlicher Teilbewegungen umfassen und drückt sich in ihrer zeitlichen Dauer und in der Bewegungsfrequenz aus (Meinel et al., 2015, S. 137). Während ein das Tempo steigernder Anlauf mit Erreichen einer hohen Endgeschwindigkeit als optimal (Qualitätsstufe 1 – ideale Ausführung) gelten kann, ist das Angehen mit stets gleichbleibendem Tempo als defizitäre Ausführung (Qualitätsstufe 3 – starke Abweichung) zu bewerten. Ein ideal ausgeführter, flacher Impulsschritt bewirkt dabei eine leichte Geschwindigkeitszunahme oder zumindest ein Halten der Geschwindigkeit zum Abwurf hin.

Der **Bewegungsrhythmus** umfasst allgemein die spezifische, charakteristische zeitliche Ordnung von Teilprozessen, die in einem motorischen Akt miteinander verbunden sind (Meinel et al., 2015, S. 92). Als Beispiel kann der Anlaufrhythmus bei der Ausführung des Impuls- und Stemmschrittes der Wurfbewegung dienen. Während die vorhergehenden Anlaufschritte noch zyklisch und durch die Temposteigerung mit zunehmender Frequenz ausgeführt werden, zeichnen sich die letzten drei Bodenkontakte durch eine spezifische zeitliche Relation aus. Durch den langgezogenen Impulsschritt entsteht eine längere Flugphase des vorletzten Schrittes, während die letzten beiden Bodenkontakte kurz nacheinander aufgesetzt werden. Somit ergibt sich der charakteristische Bewegungsrhythmus lang – kurz – kurz (oder häufig mit dem didaktischen Schlagwort Am-ster-dam bezeichnet). Wenn dieser Rhythmus äußerlich gut nachvollzogen werden kann, ist die Qualitätsstufe 1 zu wählen, während auf Qualitätsstufe 3 zu entscheiden ist, wenn die beiden letzten Schritte sich in rhythmischer Hinsicht nicht von den vorhergehenden unterscheiden.

Der **Bewegungsumfang der Ballrückführung in der Wurfauslage** ist ein weiteres, gut beobachtbares Bewegungsmerkmal, welches die räumliche Ausdehnung der Wurfbewegung beschreibt. Allerdings ist die Größe des Bewegungsumfangs ohne technische Hilfsmittel für den Beobachter nicht quantifizierbar, sondern muss über Beobachtung und Einschätzung im Vergleich zu einem Bewegungsleitbild gewonnen werden (Meinel et al., 2015, S. 134f.). Da Wurfbewegungen einen großen Gesamtimpuls erfordern, ist die Maximierung des Beschleunigungswegs durch ein möglichst weites Nach-hinten-Führen der Wurfhand zu realisieren. Dies erfordert wiederum eine möglichst ausgeprägte Oberkörperrotation nach hinten mit einer zeitgleichen Streckung des Wurfarms, so dass die Schulterachse und der gestreckte Wurfarm parallel zur Wurfrichtung nach hinten zeigen (entspricht der Qualitätsstufe 1). Diese Körperhaltung in der Wurfauslage stellt eine Schlüsselposition der Wurfbewegung dar, welche aber schwer zu beobachten ist, weil sie als Umkehrpunkt nur kurz gehalten wird. Wird hingegen der Oberkörper überhaupt nicht zurückgeführt, sondern bleibt frontal in Wurfrichtung stehen, so führt dies zusammen mit einem gebeugten Wurfarm zu einer drastischen Verkürzung des Beschleunigungswegs und somit zu einer geringeren Wurfweite (Qualitätsstufe 3).

Die **Bewegungskopplung der Teilbewegungen in der Wurfbewegung** ist ein ebenfalls schwer fassbares Ereignis, welches dem ungeschulten Auge meist nur in Zeitlupe zugänglich ist. Nach dem kurzzeitigen Einnehmen der Wurfauslage durch eine Verwindung von Hüft- und Schulterachse ist die Beschleunigung in Wurfrichtung „zuerst im Rumpf, danach in weiterer distaler Verschiebung im Schultergelenk, Ellbogengelenk und Handgelenk“ (Meinel et al., 2015, S. 109) erkennbar (Qualitätsstufe 1). Es kommt also zu einer nach außen verlaufenden Verschiebung des (Haupt-)Phasenbeginns, weil sich der Rumpf bereits in der Hauptphase befindet, während sich der Arm noch in der Vorbereitungsphase befindet oder im Schultergelenk schon die Hauptbewegung des Oberarms beginnt, während im Ellbogengelenk noch ein Beugen als Ausholbewegung für den nachfolgenden Unterarmschlag erfolgt. Dem Rumpf kommt dabei eine Sonderrolle zu, weil dieser als Bindeglied zwischen unteren und oberen Extremitäten erst ermöglicht, die Kraftwirkung der Beine auf die Arme und von da aus weiter auf den zu bewegendem Gegenstand zu übertragen. Das Einnehmen einer Bogenspannung ist ein erkennbares Zeichen dafür, dass die Bewegungskopplung der Teilbewegungen richtig abgestimmt war. Indem der Rumpf aus der Wurfauslage nach vorne beschleunigt, der Wurfarm aber gleichzeitig noch zurückgehalten wird, wölbt sich der Rumpf nach vorne und die vertikal verlaufenden Muskelzüge sind gedehnt und damit gespannt. Auch die Hüftmuskulatur nach unten und die Schultermuskulatur nach oben sind in diese Vorspannung miteinbezogen. Wenn die gespannten Muskeln dann aus der Bogenspannung heraus kontrahieren, kommt es infolgedessen zur Beugung des Rumpfes nach vorne. Ist bei einem Werfer keine Bogenspannung erkennbar, weil Hüfte, Schulter und Wurfarm nahezu synchron nach vorne beschleunigt werden, so rechtfertigt dies die Qualitätsstufe 3.

Die **Bewegungspräzision** wird im Rahmen des obigen Beurteilungsrasters im Meinel'schen Sinne als Ablaufgenauigkeit verstanden (Meinel et al., 2015, S. 127). Die räumlich-zeitlichen und dynamischen Charakteristika sollen so im Bewegungsverlauf der Wurfbewegung realisiert werden, dass das Abfangen der Bewegungsenergie mit möglichst wenig Raumverlust verbunden ist, um die Abwurfmarkierung nicht zu übertreten. Dies sollte auch bei hohem Bewegungstempo und großem Bewegungsumfang möglich sein, wohingegen bei Anfängern oft zu beobachten ist, dass gerade diese beiden Merkmale reduziert werden, um die Bewegungspräzision zu sichern. Bei Würfeln kann sich dies in einer geringen Anlaufgeschwindigkeit, einer langsamen Abwurfbewegung oder auch einer entsprechend kleinräumig angelegten Rumpf- und Wurfarmbewegung zeigen. Wenn es der Sportler schafft, das Abbremsen der Wurfbewegung direkt im Stemmschritt oder mit maximal einem Abfangschritt nach vorne zu realisieren, verdient dies die Qualitätsstufe 1. Hierbei ist jedoch zu erwähnen, dass der Abfangschritt erst erfolgen sollte, wenn die Abwurfbewegung schon beendet ist und das Wurfgerät die Hand verlassen hat. Das Halten des gestreckten Stemmschritts von seinem Aufsatz bis zum Verlassen des Wurfgeräts von der Hand ist hierbei ein wesentliches Beobachtungsmerkmal für die Fähigkeit zur Bewegungspräzision. Wenn der Abwurf überlaufen wird und es zu keinem kontrollierten Abstoppen kommt, entspricht dies der Qualitätsstufe 3.

Die **Bewegungsstärke** scheint zwar ein sehr passendes Bewegungsmerkmal für Wurfbewegungen zu sein, dient sie doch in der Lehr- und Übungspraxis schon lange als Maßbezeichnung des Krafteinsatzes, wenn zum Beispiel von der Schärfe eines Wurfes die Rede ist. In der Leichtathletik geht es hier eindeutig um die Maximierung der Bewegungsstärke, da abgesehen vom Landen des Wurfgeräts im Wurfsektor keine bessere Treffgenauigkeit erreicht werden muss. Daher sollte auch der Ballwurf möglichst kräftig und scharf ausgeführt werden. Dabei hängt die Wirkung des Muskelkrafteinsatzes aber nicht nur von seiner Größe, sondern auch von seiner Richtung und von der zeitlichen Einordnung in den Gesamtablauf der Bewegungshandlung ab. Wenn diese drei Komponenten nicht optimal aufeinander abgestimmt sind, verpufft ein Teil des Krafteinsatzes wirkungslos (Meinel et al., 2015, S. 137f.). Da aber

der Beurteilungstest auf Filmaufnahmen basiert, die nur die Wurfbewegung zeigen, nicht aber die Flugbahn und Wurfweite des Balles, würde der Einschätzung dieses Bewegungsmerkmals ein wesentlicher Bezugspunkt fehlen. Da die Bewegungsstärke außerdem von den vorigen Bewegungsmerkmalen Umfang und Kopplung bedingt wird, wird auf eine eigene Beobachtungskategorie für die Bewegungsstärke verzichtet.

4.5.2 Konzeption eines Beurteilungsrasters

Aufgrund der Überzeugung, dass die Meinel'schen Bewegungsmerkmale eine sinnvolle Orientierungshilfe für eine morphologische Bewegungsanalyse sind, wurde ein Beurteilungsraster aus sechs Bewegungsmerkmalen und drei Qualitätsstufen entwickelt (siehe AppendixTab. 53). In der linken, blauen Spalte sind die sechs zu beurteilenden Bewegungsmerkmale aufgeführt, in den grünen Spalten die jeweilige Ausführungsqualität in drei Stufen. Die gezeigten Bewegungen müssen hinsichtlich verschiedener Bewegungsmerkmale qualitativ beurteilt werden, indem entschieden werden muss, mit welcher Bewegungsqualität das jeweilige Merkmal ausgeführt wurde. Um diese Einschätzung quantitativ messbar zu machen, werden drei Qualitätsstufen entwickelt, welche den Könnensstand einer Schülerin oder eines Schülers für jedes Bewegungsmerkmal spezifisch beschreiben. Qualitätsstufe 1 bedeutet, dass das Bewegungsmerkmal ideal ausgeführt wurde und dem Sollwert voll entspricht. Bei Qualitätsstufe 2 kommt es zu geringen Abweichungen hinsichtlich des Bewegungsmerkmals, welches aber in Grundzügen erfüllt wird. Eine Bewegungsausführung auf Qualitätsstufe 3 zeigt starke Abweichungen vom Sollwert, so dass das jeweilige Bewegungsmerkmal kaum noch erfüllt wird. In dem Raster ist außerdem verbal beschrieben, wie das jeweilige Bewegungsmerkmal ausgeführt sein soll, damit man es als ideale Ausführung, geringe oder starke Abweichung entsprechend qualifizieren kann.

Die Erstellung der Kriterien für die Qualitätsstufe 1 orientiert sich am technischen Anforderungsprofil für das Grundlagentraining, da diese Kategorie der Altersstufe der gefilmten Schülerinnen und Schüler am besten entspricht. Viele Aspekte und Details des technischen Leitbilds eines erwachsenen Speerwerfers werden somit für den Beurteilungstest relativiert. Trotzdem können die Technikkriterien des Schlagballwurfs aus der Speerwurftechnik abgeleitet werden, zumal sie als erste Teilziele im Ausbildungsprozess eines jungen Speerwerfers unerlässlich sind.

Bauersfeld und Schröter (2016, S. 602) formulieren eine erste Könnensstufe, die auch für die Altersgruppe der Bewegungsbeispiele im folgenden Beurteilungstest passend erscheint:

Im ersten Jahr des Grundlagentrainings erfolgt die *Ausbildung der Abwurfbewegung* (Schlagwurf) sowie der Abwurf aus drei Schritten unter hauptsächlichlicher Nutzung des Schlagballwurfs. Gleichzeitig wird mithilfe von Imitationsübungen das Erlernen des Fünf-Schritt-Rhythmus vorbereitet.

Konkrete technische Anforderungen für die Disziplin Schlagballwurf werden in leicht modifizierter Form aus den Ausführungen von Bauersfeld und Schröter (2016, S. 603) auch für den folgenden Beurteilungstest übernommen:

Anlaufphase

- Erzeugung einer optimalen horizontalen Anlaufgeschwindigkeit
- Azyklischer Anlaufabschnitt (fünf Schritte) mit Geschwindigkeitssteigerung, jedoch noch nicht kontinuierlich (in der Regel mit Abfall bei Rückführung und Impulsschritt)

- Übergang zum azyklischen Abschnitt harmonisch ohne Trippeln oder Ziehen der Schritte
- Zeitlich richtiger Beginn der Rückführung und weitgehender Abschluss nach dem zweiten Schritt des Fünf-Schritt-Rhythmus
- Rumpfposition nach der Rückführung gekennzeichnet durch gestreckten Wurfarm, erkennbare Körperrücklage, Schulterdrehung um 90 Grad sowie leichte Hüftdrehung
- Weitgehendes Beibehalten der Rumpfrücklage bis hin zum Abwurf
- Impulsschritt als verlängerter flacher Schritt erkennbar, an dessen Ende die optimale Körperposition beim Setzen des Druckbeins deutlich ausgeprägt ist (sichtbare Beinbeugung, Schulter- und Armhaltung hinten, Hüftdrehung gegen 45°)

Abwurfphase

- Flache Streckbewegung des Stemmbeins nach vorne
- Flaches Aufsetzen des Stemmbeins, minimal gebeugt, jedoch im Kniegelenk fixiert
- Erkennbares Nacheinander im Einsatz von Bein, Hüfte, Schulter, Wurfarm
- Wurfarm bis kurz vor Stemmbeinaufsatz gestreckt
- Bogenspannung erkennbar durch Verzögerung im Einsatz von Rumpf und Arm
- Sichtbares Stemmen des linken Beins bei Tendenz der Streckung im Moment des Abwurfs
- Pronation des rechten Beins beim Abwurf
- Erkennbare Schlagbewegung des Unterarms

In Tab. 53 des Appendix sind die Qualitätsstufen für die jeweiligen Bewegungsmerkmale in Stichpunkten ausformuliert. Dieses Beurteilungsraster liegt den Testpersonen beim Beurteilungstest in ausgedruckter Form vor.

4.5.3 Beurteilungsaufgaben

Als Bewegungsbeispiele für den Beurteilungstest dienten Filmaufnahmen einer Schülerleichtathletikgruppe des 1. FC Passau (gefilmt im Zeitraum Juli 2015 bis Juli 2016) und einer siebten Jahrgangsstufe (Jungen und Mädchen) des Gymnasiums Untergriesbach (gefilmt im Zeitraum März und April 2017). Die Schülerinnen und Schüler wurden angewiesen, einen Schlagwurf mit fünf Schritten Anlauf auszuführen. Es wurde nicht vorgegeben, ob das Wurfgerät schon zurückgeführt sein sollte oder erst im Laufe des Anlaufs zurückgeführt wird, weil Anfänger den zyklischen Anlaufteil ohnehin oft ganz weglassen und mit bereits zurückgeführtem Wurfgerät in seitlicher Ausgangsposition beginnen (Strüder et al., 2016, S. 669). Es erfolgte jedoch der Hinweis, dass die Abwurflinie nach dem Wurf nicht übertreten werden sollte.

Die fünf Teile des Beurteilungstests wurden grundsätzlich so konzipiert, dass die Komplexität bei jeder neuen Aufgabenstellung steigt. Jede Bewegungsausführung wird nur ein einziges Mal gezeigt, ehe die Beurteilung des Bewegungsmerkmals erfolgen muss. Den Probanden wird vor Beginn des Beurteilungstests mittels Videobeispielen genau erklärt, wie die Bewegungsmerkmale in der Altersstufe des Grundagentrainings idealerweise ausgeführt sein sollten. Auch die Qualitätsstufen 2 und 3 werden mit Bewegungsbeispielen konkretisiert, so dass die Probanden eine möglichst einheitliche Einschätzung hinsichtlich der Qualitätsstufen treffen können sollten. Auf diese Weise soll sichergestellt werden, dass die Testpersonen unter den abstrakten Bewegungsmerkmalen das gleiche verstehen und auch die Qualitätsstufen in gleicher Weise zueinander abgrenzen.

Teil A des Beurteilungstests besteht dann aus sechs Zuordnungsaufgaben (für jedes Bewegungsmerkmal eine Aufgabe), bei der die Testperson pro Aufgabe drei verschiedene Bewegungsausführungen hintereinander in Echtzeit gezeigt bekommt. Sie muss dann alle drei Bewegung der passenden Qualitätsstufe des vorgegebenen Bewegungsmerkmals zuordnen, wobei jede Qualitätsstufe nur einmal gewählt werden darf. Als methodische Hilfe steht das Beurteilungsraster (siehe Appendix Tab. 53) zur Verfügung, in welchem die wichtigsten Kriterien der jeweiligen Qualitätsstufen nochmals zusammengefasst sind. Teil A ist somit nicht nur der erste Testteil, sondern soll auch das Verständnis der Bewegungsmerkmale und ihrer Qualitätsstufen vertiefen, so dass alle Testpersonen für die immer komplexer werdenden Beurteilungsaufgaben vergleichbare Voraussetzungen haben.

Teil B besteht aus zwölf Aufgaben, wobei für jede Bewegungsausführung die Qualitätsstufe eines vorgegebenen Bewegungsmerkmals bestimmt werden muss. Die Komplexitätssteigerung zu Teil A besteht darin, dass die Testperson zwar das Bewegungsmerkmal kennt, das sie beurteilen soll, jedoch entscheiden muss, welche Qualitätsstufe sie als passend erachtet, ohne dies mit anderen Bewegungsbeispielen vergleichen zu können. Die Reihenfolge der Bewegungsmerkmale, welche jeweils zwei Aufgaben umfassen, bleibt unverändert: Bewegungsharmonie, Bewegungstempo, Bewegungsrhythmus, Bewegungsumfang, Bewegungskopplung und Bewegungspräzision.

Teil C ist als Bewegungsvergleich konzipiert, bei dem pro Aufgabe wieder ein Bewegungsmerkmal vorgegeben ist und zwei unterschiedliche Bewegungsausführungen des gleichen Athleten direkt hintereinander gezeigt werden. Die Testperson muss entscheiden, ob das Bewegungsmerkmal bei der zweiten Ausführung besser, gleich oder schlechter ausgeführt wurde als bei der ersten Ausführung. Für das Merkmal Bewegungstempo werden die Begriffe schneller, gleich oder langsamer und für den Bewegungsumfang die Begriffe größer, gleich oder kleiner verwendet. Die Komplexitätssteigerung besteht darin, dass die beiden Bewegungsausführungen zunächst für sich beurteilt und dann miteinander verglichen werden müssen. Da es sich auch noch um den gleichen Sportler handelt und somit der gesamte Bewegungsablauf recht ähnlich ist, muss das Bewegungsmerkmal schon sehr genau beobachtet werden können, um eine Entscheidung zu treffen. Die Aufgabe ist der realitätsnahen Beurteilungssituation nachempfunden, dass eine Sportlehrkraft immer wieder Bewegungsmerkmale miteinander vergleichen können muss, wenn sie einem Schüler eine Rückmeldung geben will, ob sich eine Verbesserung oder Verschlechterung im Vergleich zu vorhergehenden Bewegungsausführungen ergeben hat.

Teil D besteht aus zwei Aufgaben und kann als Fehleranalyse oder Fehleridentifikationsaufgabe bezeichnet werden. Die Testperson bekommt pro Aufgabe ein Bewegungsbeispiel zu sehen, für das sie zum einen die Qualitätsstufe der Bewegungsharmonie bestimmen muss, zum anderen alle Bewegungsmerkmale identifizieren soll, welche der Qualitätsstufe 3 entsprechen, also als stark abweichend von der idealen Bewegungsausführung eingeschätzt werden. Diese Beurteilungsaufgabe entspricht dem methodischen Grundsatz der Fehlerkorrektur, zuerst die Hauptfehler bzw. schwerwiegenden Bewegungsabweichungen zu korrigieren, ehe auf geringere Abweichungen geachtet wird. Zu diesem Zwecke muss die Lehrkraft in der Lage sein, stark abweichende Bewegungsmerkmale zu identifizieren und dem Schüler darüber eine Rückmeldung zu geben.

Teil E verlangt anhand von zwei Beispielen eine Gesamtanalyse aller Bewegungsmerkmale bei nur einem Beobachtungsdurchgang in Echtzeit. Die Testperson muss für alle sechs Bewegungsmerkmale eine Einschätzung der jeweiligen Qualitätsstufe treffen. Es ist zu erwarten, dass diese Beurteilungsaufgabe am schwersten zu lösen ist, weil sich die Testperson nicht auf alle Bewegungsmerkmale gleichzeitig konzentrieren kann und daher einige Einschätzungen misslingen könnten.

4.5.4 Inter-Coder-Reliabilität

Vor der Anwendung des Tests zur Bewegungsbeurteilung wurden die einzelnen Testteile von fünf ausgewiesenen Experten, die das Bewegungssehen und Bewegungsbeurteilen in der Leichtathletik regelmäßig auf hohem Niveau praktizieren, eingeschätzt. Als Experten standen Fleischmann Roland, Lehner Christoph, Lehner Karl, Bloch Florian und Reuter Andreas zur Verfügung, deren Expertise in Tab. 18 aufgelistet ist. Jeder Experte durchlief den Beurteilungstest wie ein Proband und gab anschließend ein Urteil über die gezeigten Bewegungsbeispiele ab. Zu sehen bekamen die Experten die Bewegungsbeispiele zunächst nur einmal, um herauszufinden, ob die Beurteilungsaufgaben von erfahrenen Beobachtern in einstimmiger Weise gelöst werden oder ob es Uneinigkeit zwischen den Experten gibt, welche auf die inhaltliche Mehrdeutigkeit eines Bewegungsbeispiels hindeutet. Nach dem ersten Urteil konnten die Experten das Video so oft wie nötig wiederholen und ihre Einschätzung gegebenenfalls revidieren. Für die weitere Auswertung wurde dasjenige Urteil verwendet, welches sich auch nach mehrfacher Betrachtung in den Augen des jeweiligen Experten als stabil erwies.

Tab. 18: *Rater des Bewegungsbeurteilungstests mit Qualifikationsnachweis*

Fleischmann Roland	BLV-Leichtathletik-Trainer
Lehner Christoph	B-Trainer Leichtathletik, Sportlehrkraft an Gymnasien, BLV-Trainertätigkeit
Lehner Karl	C-Trainer Leichtathletik, Sportlehrkraft an Grundschulen
Bloch Florian	Dozent in der Sportlehrerausbildung an der Universität Würzburg mit Lehr- und Prüfertätigkeit in Leichtathletik
Reuter Andreas	Dozent in der Sportlehrerausbildung an der Universität Würzburg mit Lehr- und Prüfertätigkeit in Leichtathletik

Zur allgemeinen Überprüfung der Einheitlichkeit der Expertenurteile wird die sog. Inter-Coder-Reliabilität verwendet. Sie gibt das Ausmaß an, in dem mehrere unabhängige Kodierer (Rater/Experten) mit derselben Kodiervorschrift zur selben Einschätzung einer Beurteilungsaufgabe kommen. Zu ihrer Überprüfung wird die Berechnung von Krippendorff's Alpha (r_α) gewählt, welche das konservativste Verfahren zur Reliabilitätsbestimmung bei mehreren Ratern darstellt und bezüglich anderer Verfahren wie dem Übereinstimmungskoeffizienten, Scotts Pi oder Cohens Alpha den Vorteil hat, auch bei mehreren Expertenurteilen und unvollständigen Datensätzen (welche in dieser Studie nicht vorlagen) angewendet werden zu können sowie unabhängig von der Stichprobengröße zu sein (Hayes & Krippendorff, 2007). Die Berechnung erfolgt mittels einer zusätzlichen SPSS-Makrodatei, welche eine Syntaxerweiterung mit der Bezeichnung Kalpha in SPSS anlegt (Swert, 2012).

Für Teil A des Beurteilungstests (Einschätzung der Bewegungsmerkmale) ergibt sich als Inter-Coder-Reliabilität zwischen den fünf Ratern Fleischmann, Lehner C., Lehner K., Bloch und Reuter ein zufriedenstellender Wert von $r_\alpha = 0,8184$. Zwar werden von einigen Autoren Werte über $r_\alpha = 0,9$ (als gut) bzw. über $r_\alpha = 0,8$ (als angemessen), jedoch geben sich viele Autoren für explorative Untersuchungen (wie dieser Studie) durchaus auch mit einer Reliabilität von $r_\alpha = 0,7$ oder geringer zufrieden (Neuen-

dorf, 2010, S. 145). Die Bewegungsbeispiele zu den Merkmalen Bewegungsharmonie (A1), Bewegungsumfang (A4) und Bewegungspräzision (A6) wurden von den fünf Experten ohne Abweichung übereinstimmend beurteilt. Das Merkmal Bewegungstempo wurde uneinheitlich bewertet (siehe Tab. 19), wobei bei Item A2.1 drei Experten und bei den Items A2.2 und A2.3 vier Experten in ihrem Urteil übereinstimmten, so dass auch für dieses Item ein Hauptexpertenurteil (A2.1 als QS3, A2.2 als QS2 und A2.3 als QS1) gebildet werden kann. Beim Bewegungsrhythmus waren sich die Experten bei zwei Items (A3.1 und A3.3) vollständig einig, bei A3.2 weicht lediglich Rater C von den anderen vier geringfügig ab. Ebenso gab es bei der Bewegungskopplung völlige Übereinstimmung in zwei Items (A5.2 und A5.3) und wiederum eine geringfügige Abweichung bei A5.1 durch Rater C.

Tab. 19: *Expertenurteile für Item A2*

A2: Bewegungstempo		obsa	obsb	obsc	obsd	obse
A2.1	3	2	2	3	3	
A2.2	2	1	2	2	2	
A2.3	1	1	2	1	1	

Es bleibt also festzuhalten, dass das Bewegungstempo von den Experten als eher schwer objektiv beurteilbar betrachtet wird, wohingegen die Beispielbewegungen der anderen fünf Merkmale vollständig bis relativ eindeutig beurteilt werden können. Würde man die Items des Merkmals Bewegungstempo eliminieren, würde man sogar einen Wert von $r_\alpha = 0,9211$ erreichen. Auf diese Eliminierung wird jedoch verzichtet, da das Merkmal Bewegungstempo trotz dieser Schwierigkeiten der Experteneinschätzung für die morphologische Bewegungsanalyse als wichtig erachtet wird.

Aus Teil B werden die ursprünglichen Items B6 und B11 entfernt, weil hier die Abweichung zwischen mindestens drei einigen Experten und mindestens einem abweichenden Experten größer als nur eine Qualitätsstufe ist und dies auf eine gravierende Mehrdeutigkeit des Videobeispiels im Bezug auf das verlangte Bewegungsmerkmal hindeutet.³⁷ Durch diese Eliminierung wird für die 10 Items ein alpha-Wert von $r_\alpha = 0,6419$ erreicht. Eine einstimmige Übereinkunft zwischen den fünf Ratern kommt bei den Items B1 und B2 (Bewegungsharmonie), B3 (Bewegungstempo), B7 (Bewegungsumfang) sowie B9 und B10 (Bewegungskopplung) vor. Bei den restlichen Items B4, B5, B8 und B12 sind es jeweils zwei, Experten, deren Abweichung allerdings auf der gleichen Qualitätsstufe liegt (siehe Tab. 20). Die Bewegungsbeispiele der Merkmale Rhythmus und Präzision werden am uneinheitlichsten von den Experten bewertet, denn in keinem der beiden Items wurde eine vollständige Expertenübereinkunft erzielt.

³⁷ Krippendorff's Alpha erreicht für alle Items (B1 bis B12) einen Wert von $r_\alpha = 0,5129$.

Tab. 20: Expertenurteile für Teil B des Beurteilungstests

		Obsa	obsb	obsc	obsd	Obse
Bewegungsharmonie						
	B1	3	3	3	3	3
	B2	2	2	2	2	2
Bewegungstempo						
	B3	1	1	1	1	1
	B4	2	3	3	2	2
Bewegungsrhythmus						
	B5	1	1	1	2	2
	B6	3	2	1	3	3
Bewegungsumfang						
	B7	1	1	1	1	1
	B8	3	3	3	2	2
Bewegungskopplung						
	B9	3	3	3	3	3
	B10	2	2	2	2	2
Bewegungspräzision						
	B11	1	3	3	2	2
	B12	3	2	2	2	3

Teil C des Beurteilungstests besteht aus sechs Beispielvideos, in denen zwei Wurfbewegungen desselben Sportlers miteinander verglichen werden müssen. Die Probanden müssen einschätzen, ob ein vorgegebenes Bewegungsmerkmal verbessert, gleich oder verschlechtert (bzw. beim Bewegungsumfang größer, gleich oder kleiner sowie beim Bewegungstempo schneller, gleich oder langsamer) ausgeführt wird. Bei den Beispielen C1 bis C6 dient ein einzelnes Bewegungsmerkmal als Vergleichskriterium, bei den Beispielen C7 und C8 sind es zwei Bewegungsmerkmale. Für alle sechs Items beträgt Krippendorfs Alpha $r_\alpha = ,0,6051$. Dabei wurden die Beispiele C1 (Bewegungsrhythmus), C3 (Bewegungskopplung) und C5 (Bewegungspräzision) von allen fünf Experten mit der gleichen Einschätzung beurteilt. Bei den Beispielen C2 (Bewegungsharmonie) und C6 (Bewegungstempo) weicht lediglich ein Experte um eine benachbarte Kategorie vom Hauptexpertenurteil ab, beim Beispiel C4 kommt es zu zwei abweichenden Urteilen. Es kann also aufgrund der Expertenübereinstimmung ein relativ plausibles Hauptexpertenurteil eruiert werden, weil sich bei jedem Item mindestens drei Experten in ihrer Einschätzung einig sind und die übrigen Experten höchstens um eine Qualitätsstufe davon abweichen (siehe Tab. 21).

Tab. 21: *Expertenurteile für Teil C des Beurteilungstests*

	Obsa	obsb	obsc	obsd	obse
C1 Bewegungsrhythmus	1	1	1	1	1
C2 Bewegungsharmonie	2	1	1	1	1
C3 Bewegungskopplung	3	3	3	3	3
C4 Bewegungsumfang	2	1	1	1	2
C5 Bewegungspräzision	1	1	1	1	1
C6 Bewegungstempo	3	3	3	3	2
C7.1 Bewegungsumfang	1	2	2	1	2
C7.2 Bewegungsrhythmus	1	1	1	1	2
C8.1 Bewegungstempo	2	3	3	3	3
C8.2 Bewegungspräzision	2	2	1	2	2

In Teil D geht es um die Aufgabe, bei einer Wurfbewegung die Bewegungsharmonie mittels der Qualitätsstufen 1 (= ideale Bewegungsausführung), 2 (= geringe Abweichung vom Sollwert) und 3 (= starke Abweichung vom Sollwert) einzuschätzen und gleichzeitig zu entscheiden, welches der fünf Bewegungsmerkmale (Rhythmus, Tempo, Umfang, Kopplung und Präzision) so defizitär ausgeführt wurde, dass es mit Qualitätsstufe 3 zu bewerten ist. Bezüglich der Interrater-Reliabilität herrscht bei diesem Testteil mit einem Wert für Krippendorff's Alpha von $r_\alpha = 0,6382$ eine gute Übereinstimmung zwischen den fünf Experten bei der Beurteilung der sechs Bewegungsbeispiele. In der graphischen Darstellung der Einzelitems ist erkennbar, dass auch für die Aufgaben D1 und D2 die Mindestanforderung der Expertenübereinkunft erfüllt ist, dass mindestens drei Experten einer Meinung sind bezüglich ihres Urteils und im Fall der Einschätzung der Bewegungsharmonie die übrigen Experten die benachbarte Qualitätsstufe gemeinsam belegen (siehe Tab. 22).

Tab. 22: *Expertenurteile für die Aufgaben D1 und D2 des Beurteilungstests*

	obsa	obsb	obsc	obsd	obse
D1 Bewegungsharmonie	3	3	3	3	3
D1.1 Bewegungstempo	0	1	0	1	0
D1.2 Bewegungsrhythmus	1	1	1	0	1
D1.3 Bewegungsumfang	1	1	1	1	1
D1.4 Bewegungskopplung	1	1	1	1	1
D1.5 Bewegungspräzision	0	0	0	1	0
D2 Bewegungsharmonie	3	2	3	2	3
D2.1 Bewegungstempo	0	0	0	0	0
D2.2 Bewegungsrhythmus	1	0	0	0	1
D2.3 Bewegungsumfang	1	1	1	1	1
D2.4 Bewegungskopplung	1	1	1	1	1
D2.5 Bewegungspräzision	1	1	1	1	1

Teil E ist die anspruchsvollste Aufgabe des Beurteilungstests. Die Probanden müssen die Qualität aller sechs Bewegungsmerkmale nach nur einem Beobachtungsdurchgang qualitativ einschätzen. Für die Teile E1 und E2 liegt der alpha-Wert lediglich bei $\alpha = 0,2373$. In Tab. 23 lässt sich erkennen, dass es bei E1 und E2 nur zwei Items (E1.4 und E2.6) gibt, in denen die Experten einstimmig beurteilen. Offensichtlich ist es sogar bei Experten kaum möglich, trotz mehrmaliger Wiederholung der Bewegungsbeobachtung übereinstimmende Einschätzungen der sechs Bewegungsmerkmale zu erreichen. Trotzdem trifft die eben erwähnte Mindestanforderung bezüglich der Übereinstimmungsqualität der Expertenurteile bei allen Items außer E1.5 zu (siehe Tab. 23). Es ist daher durchaus möglich, ein plausibles Hauptexpertenurteil zu bestimmen. Aufgrund der häufigen Uneinigkeiten erscheint es jedoch angezeigt, der Auswertungsmethode des ordinalen Rasch-Modells den Vorzug gegenüber dem dichotomen Rasch-Modell zu geben (siehe Kapitel 4.7.2.2), da für letzteres eine Einteilung der Antworten in richtig oder falsch bei vielen Items zumindest fragwürdig erscheint.

Tab. 23: Expertenurteile für die Aufgaben E1 und E2 des Beurteilungstests

		obsa	obsb	obsc	obsd	obse
E1	E1.1	2	2	2	3	2
	E1.2	2	2	2	3	2
	E1.3	2	1	1	2	2
	E1.4	2	2	2	2	2
	E1.5	2	2	1	3	2
	E1.6	1	2	2	2	1
E2	E2.1	3	3	2	3	3
	E2.2	3	2	2	3	3
	E2.3	3	2	2	3	3
	E2.4	3	3	2	3	3
	E2.5	3	3	2	2	3
	E2.6	3	3	3	3	3

4.5.5 Gütekriterien

4.5.5.1 Objektivität

Um den Beurteilungstest so objektiv wie möglich durchführen zu können, wird der Ablauf anhand eines Testmanuals immer in der gleichen Weise absolviert. Der Testraum ZMKU10 wird so vorbereitet, dass fünf Arbeitsplätze zentral vor einem Smartboard-Bildschirm und mit jeweils drei Metern Abstand zu diesem besetzt werden können. Jeweils fünf Studierende nehmen pro Testtermin teil, um in ungestörter und konzentrierter Form die Aufgaben bewältigen zu können. Vor Beginn des Tests wird den Studierenden ein Platz zugewiesen, der schon mit den benötigten Materialien (Beschreibung der Qualitätsstufen, Antwortbogen) vorbereitet ist. Vor Testbeginn erklärt der Testleiter den Testpersonen den Zweck dieser Untersuchung und bespricht das Beurteilungsraster mit den sechs ausgewählten Meinel'schen Bewegungsmerkmalen sowie den drei Qualitätsstufen. Es erfolgt außerdem der Hinweis, dass die zu bewertenden Bewegungen von Schülerinnen und Schülern einer siebten Jahrgangsstufe ausgeführt wurden.

Vor Beginn eines Aufgabenteils wird die konkrete Aufgabenstellung vom Versuchsleiter aus dem Testmanual verlesen. Die Vorgabe für Aufgabe A lautete beispielsweise:

Zuerst bekommen Sie das jeweilige Bewegungsmerkmal, das Sie bewerten müssen, mithilfe eines Kurzvideos erklärt, welches auch Beispiele für die einzelnen Qualitätsstufen enthält. Anschließend müssen drei Bewegungsausführungen der jeweiligen Qualitätsstufe zugewiesen werden. Pro Qualitätsstufe gibt es eine Bewegung, das heißt es gibt keine zwei Beispiele auf der gleichen Qualitätsstufe. Wenn Sie zwei Bewegungsausführungen ähnlich einschätzen, müssen Sie sich trotzdem für zwei unterschiedliche Qualitätsstufen entscheiden. Alle drei Bewegungsausführungen bekommen Sie nur einmal in Originalgeschwindigkeit gezeigt. Ihre Einschätzung kreuzen Sie dann auf dem Auswertungsbogen an.

Aufgrund der geringen Probandenzahl pro Testtermin konnte darauf geachtet werden, dass jede Aufgabenstellung bearbeitet und der Auswertungsbogen vollständig ausgefüllt wurde. Somit gab es bei der Auswertung keine Schwierigkeiten durch fehlende Werte. Die Auswertungsbögen wurden mittels Scanner digitalisiert und im .tif-Format in die Auswertungssoftware Klaus importiert. Nach der automatisierten Auswertung erfolgte die Datenausgabe in einer Exceltabelle, in welcher alle gekreuzten Antworten mit der Ziffer 1 und alle nicht-gekreuzten Felder mit der Ziffer 0 codiert wurden. Der Vergleich der Rohdaten mit den richtigen Antworten erfolgte dann mittels Excel-Befehl WENN (Prüfung; Dann_Wert; Sonst_Wert), so dass nach dieser Überprüfung eine dichotomisierte Datenmatrix vorlag, in welcher festgehalten ist, ob eine Testperson eine Bewegungsausführung richtig (= Ziffer 1) oder falsch (= Ziffer 2) beurteilt hat. Durch diese größtenteils automatisierten Abläufe wird eine hohe Auswertungsobjektivität gewährleistet.

4.5.5.2 Reliabilität

Zieht man Cronbach's Alpha heran, um die Reliabilität des gesamten Beurteilungstests zu überprüfen, so kommt man bei Berücksichtigung aller 49 Items auf einen Wert von $\alpha = 0,533$ (für standardisierte Items). Wenn man nur die Teile des Beurteilungstests einzeln auf interne Konsistenz überprüft, so zeigen sich sehr unterschiedliche Ergebnisse (siehe Tab. 24).

Tab. 24: Cronbach's Alpha für die Einzelteile des Beurteilungstests

	Cronbach's Alpha	Cronbach's Alpha für standardisierte Items	Guttman's Split-Half-Koeffizient
Teil A	,562	,601	,588
Teil B	-,102	-,129	,032
Teil C	,111	,055	,155
Teil D	,097	,087	-,088
Teil E	,504	,541	,386

Teil A und Teil E liefern im Gegensatz zu den Teilen B, C und D zufriedenstellende Ergebnisse bezüglich der internen Konsistenz. Jedoch muss man bedenken, dass diese Berechnung ganz ohne Berücksichtigung der sechs Bewegungsmerkmale erfolgt. Möglicherweise hängt es auch vom Bewegungsmerkmal ab, wie leicht oder schwer Testpersonen die Beurteilungsaufgabe fällt. Bei obigen Berechnungen wird lediglich davon ausgegangen, dass die Aufgabenstellung die gleiche ist. Fasst man hingegen alle Items

entsprechend des jeweiligen Bewegungsmerkmals zusammen, das eingeschätzt werden muss, so fällt anhand der Alpha-Werte auf, dass die Merkmale Harmonie, Rhythmus und Kopplung eine deutlich höhere, wenngleich immer noch unzufriedenstellende interne Konsistenz aufweisen als dies bei Bewegungstempo, Bewegungsumfang und Bewegungspräzision der Fall ist. Vor allem die Bewegungspräzision weist sogar einen hohen negativen Wert auf, was daran liegt, dass die Items zueinander nur schwach oder sogar negativ miteinander korrelieren (siehe Tab. 25). Es kann schlussgefolgert werden, dass die Merkmale Tempo, Umfang und vor allem Präzision weniger reliabel sind als die Merkmale Harmonie, Rhythmus und Kopplung.

Tab. 25: Cronbach's Alpha für die Bewegungsmerkmale des Beurteilungstests

	Cronbach's Alpha	Cronbach's Alpha für standardisierte Items	Guttman's Split-Half-Koeffizient
Bewegungsharmonie	,286	,261	-,030
Bewegungstempo	-,006	-,030	-,026
Bewegungsrhythmus	,346	,330	,418
Bewegungsumfang	,092	,023	-,007
Bewegungskopplung	,492	,535	,367
Bewegungspräzision	-,403	-,279	-,557

4.5.5.3 Validität

Die Beurteilungsleistung sportlicher Bewegungen zu operationalisieren stellt eine schwierige Aufgabe dar. Es hängt wesentlich vom Zweck der Beurteilung ab, wie genau die Bewegungsbeobachtung erfolgen muss. Wenn man an schulische Leistungsbewertung denkt, so reicht es aus, die Ausführungsqualität (im Idealfall anhand vorgegebener Kriterien) in einer von sechs Schulnoten auszudrücken. Will man hingegen einem Schüler einen Korrekturhinweis geben, wie er seine Bewegungen ausführen verbessern kann, so muss man zuerst detailliert beobachten, in welchen Phasen der Bewegung funktionale Schwächen vorliegen, bevor man sie verbalisieren und rückmelden kann. Die Beurteilung, ob ein Korrekturhinweis bei der nächsten Bewegungen ausführung zielführend vom Sportler umgesetzt werden konnte, stellt wiederum eine andere Herausforderung dar. Aufgrund dieser Tatsache erscheint es kaum gerechtfertigt, die Beurteilungsleistung sportlicher Bewegungen als eindimensionales Konstrukt zu beschreiben, sondern sich über konkrete Beurteilungssituationen der Komplexität dieses Objekts anzunähern. Dass das Beurteilen anhand von Videoaufnahmen erfolgt und nicht durch Beobachten realer Wurfbewegungen, bringt sicherlich Verzerrungen mit sich, da die zweidimensionale Darstellungsform über Videos nur begrenzt der realen Beurteilungssituation entspricht und Ungeübten das Bewegungssehen dadurch schwerer fallen könnte. Jedoch relativiert sich diese Schwäche durch den Vorteil der Vergleichbarkeit der Ergebnisse, welche nur dadurch gewährleistet werden kann, dass jeder Testperson die gleichen Testbedingungen geboten werden können.

Um die Komplexität der Beurteilungsleistung angemessen wiederzugeben, wird auch der Beurteilungstest in mehrere Teile untergliedert, welche eng an realen Beurteilungssituationen einer Sportlehrkraft oder eines Trainers orientiert sind. Grundlegend werden vier Fähigkeitsbereiche unterschieden, welche bei der Beurteilung sportlicher Bewegungen eine Rolle spielen: 1. Das Einschätzen eines einzelnen

Bewegungsmerkmals, 2. der Vergleich eines Bewegungsmerkmals anhand zweier Bewegungsausführungen, 3. die Identifikation defizitärer Bewegungsmerkmale im Sinne einer Fehleranalyse und 4. die Gesamtanalyse aller Bewegungsmerkmale.

Eine im Sport bewertende Person muss demnach in der Lage sein, die Ausführungsqualität einzelner Bewegungsmerkmale (in dieser Studie anhand von drei Qualitätsstufen) beurteilen zu können. Diese Fähigkeit wird in den Teilen A und B überprüft, indem die Testperson Bewegungsausführungen betrachtet, für welche sie nur eine Einschätzung bezüglich der Qualitätsstufe eines vorgegebenen Bewegungsmerkmals treffen muss. In Teil C wird die Beurteilungsleistung dahingehend gemessen, dass die Testperson zwei verschiedene Bewegungsausführungen des gleichen Athleten miteinander vergleichen muss, um einzuschätzen, ob ein vorgegebenes Bewegungsmerkmal besser, schlechter oder gleich gut ausgeführt wurde (Bewegungsvergleich). Teil D steigert die Beurteilungskomplexität, indem alle Bewegungsmerkmale beobachtet werden müssen, um die stark von der idealen Bewegungsausführung abweichenden Merkmale (entspricht der Qualitätsstufe 3) identifizieren zu können. Diese Fähigkeit kann vom Grundsatz der Fehleranalyse, zuerst die schwerwiegenden Hauptfehler zu korrigieren, bevor man die Bewegungsdetails in den Blick nimmt, abgeleitet werden. Die letzte Komplexitätsstufe stellt die Gesamtanalyse aller Bewegungsmerkmale in Teil E dar, welche eine sehr differenzierte Bewegungswahrnehmung der Testperson erfordert, um die zum Teil synchron ablaufenden Bewegungsmerkmale auf der richtigen Qualitätsstufe einzuordnen. Durch die enge Anlehnung dieser vier Aufgaben an reale Beurteilungssituationen sowie die Verwendung von Bewegungsbeispielen realer Schülerinnen und Schüler, die genau in dieser Weise auch im Sportunterricht beurteilt werden müssen, wird versucht, eine hohe Validität des Beurteilungstests zu erreichen.

4.6 Vorerfahrung im Beurteilen sportlicher Bewegungen

Es ist zu erwarten, dass die Vorerfahrungen im Beurteilen sportlicher Bewegungen eine Rolle spielen, wenn es darum geht, die Beurteilungsleistung eines Probanden im Beurteilungstest sinnvoll zu interpretieren. Dabei kann durchaus angezweifelt werden, dass erfahrene Bewegungsbeobachter in Sportarten, die mit der Schlagwurfbewegung nichts zu tun haben, einen Vorteil gegenüber generell unerfahrenen Bewegungsbeobachtern haben. Hingegen könnte angenommen werden, dass die Beobachtungserfahrung in Sportarten, in denen dem leichtathletischen Wurf ähnliche Schlagbewegungen zur sportlichen Handlungspraxis gehören, auch bei den in dieser Studie verwendeten Beurteilungsaufgaben hilfreich sind. Dies könnte beispielsweise beim Volleyball (Aufschlag von oben, Schmetterschlag), Badminton (Smash oder Clear), Tennis (Aufschlag) oder Handball (Schlagwurf) der Fall sein.

Darum wird mittels eines Fragebogens (siehe Abb. 119 im Appendix) die Vorerfahrung im Beurteilen sportlicher Bewegungen über eine Selbstauskunft bzw. Selbsteinschätzung der Probanden erhoben. Der Fokus ist dabei zunächst eng, indem danach gefragt wird, ob Beurteilungserfahrungen von leichtathletischen Wurfbewegungen vorliegen, die beispielsweise durch Trainerausbildung oder Trainertätigkeit erworben wurden. Die nächste Frage richtet sich auf die Beurteilung von Bewegungen, die dem leichtathletischen Schlagwurf ähnlich sind, ehe auf einer letzten Stufe erfasst wird, ob generell Vorerfahrungen in der Beurteilung anderer sportlicher Bewegungen vorliegen, die nichts mit Schlag(wurf)bewegungen zu tun haben. Das Ausfüllen des Fragebogens erfolgt vor dem Bearbeiten des Beurteilungstests.

4.7 Datenauswertung

Da die meisten Testverfahren für diese Studie selbst entwickelt werden mussten, weil es keine bereits vorhandenen Instrumente gibt, besteht ein zentrales Anliegen der Datenauswertung darin, die Güte der verwendeten Testverfahren empirisch zu prüfen. Das Testmodell soll dabei den Zusammenhang zwischen dem beobachteten Antwortverhalten der Person in verschiedenen Testaufgaben und der Ausprägung des nicht beobachtbaren (latenten) Personenmerkmals, das durch diese Aufgaben erfasst werden soll, möglichst umfassend beschreiben. Eine lineare Beziehung zwischen der Anzahl gelöster Aufgaben und der Ausprägung des Personenmerkmals anzunehmen, lässt die Tatsache unberücksichtigt, dass Aufgaben eines Leistungstests unterschiedliche Schwierigkeiten besitzen. Neumann (2014, S. 356) stellt richtigerweise fest, dass „Schlussfolgerungen über die Fähigkeit zweier Personen, die gleich viele, aber unterschiedlich schwierige Aufgaben richtig bearbeitet haben, [...] ohne weitere Berücksichtigung der Schwierigkeiten der jeweiligen Aufgaben problematisch“ sind. Dies betonen auch Bond und Fox (2015, S. 48), indem sie darauf hinweisen, dass der Rohpunktwert eines Tests allein noch keine Rückschlüsse auf die Personenfähigkeit zulässt, sondern verglichen werden muss mit dem Antwortverhalten der jeweiligen Testgruppe. Über die Einbeziehung der Aufgabenschwierigkeit kann herausgefunden werden, ob ein erreichter Rohpunktwert mit einer hohen Wahrscheinlichkeit der Personenfähigkeit entspricht oder aufgrund des individuellen Antwortverhaltens nicht zur Personenfähigkeit passt. Aufgrund dieser Überlegungen wird in dieser Studie bevorzugt das **probabilistische Rasch-Modell** als Testmodell zur Skalierung der Daten verwendet, welches die Aufgabenschwierigkeit in Beziehung zur Schätzung der Ausprägung des zu messenden latenten Merkmals setzt. Zugleich kommt es bei diesem Modell zu einer Umwandlung der ordinalskalierten Rohpunktwerte in intervallskalierte Measures, wodurch die Möglichkeiten einer späteren statistischen Auswertung erweitert werden. Bühner (2011, S. 506) betont außerdem, dass bei probabilistischen Testmodellen die Messgenauigkeit für jede Stufe der Merkmalsausprägung geschätzt werden kann, während diese bei der klassischen Testtheorie für alle Bereiche als gleich angenommen werden muss. Durch diese Tatsache ist die probabilistische Berechnung hinsichtlich der Schätzgenauigkeit der Personenparameter im Vorteil. Gleichwohl bleibt aber festzuhalten, dass die klassische und die probabilistische Testtheorie keine konkurrierenden Systeme sind, die sich gegenseitig ausschließen, sondern ihre jeweiligen Vor- und Nachteile haben, die bei der Datenauswertung berücksichtigt werden müssen. Daher wird in dieser Studie über den geeigneten Einsatz der jeweiligen Methode an geeigneter Stelle nachgedacht und entschieden, welche Theorie am plausibelsten erscheint und zum vorliegenden Datenmaterial passt. Für die probabilistischen Berechnungen wird die Software R verwendet, für alle weiteren statistischen Berechnungen die Software SPSS.

4.7.1 Die Item-Response-Theorie

Bei der Datenauswertung mittels probabilistischer Item-Response-Theorie (IRT) werden Testergebnisse nicht (wie bei der klassischen Testtheorie) als wahre (messfehlerbehaftete) Merkmalsausprägungen angesehen, sondern als Indikatoren für die zugrundeliegende latente Variable angesehen. Latente Variablen sind dabei „nicht beobachtbare, hinter den Antworten der Testperson ‚verborgene‘ Fähigkeiten, von denen das manifeste Antwortverhalten abhängig abhängig ist“ (Rentzsch & Schütz, 2009, S. 152). In einem Testverfahren stellt beispielsweise das Fachwissen die latente Variable dar, welches

den systematischen Zusammenhang zwischen den verwendeten Items bzw. gegebenen Itemantworten erklären soll. Verschiedenen Ausprägungen des Fachwissens von Testpersonen liegen dabei auf einem latenten Kontinuum und können jeden beliebigen Wert annehmen. Von der Beantwortung bestimmter Items wird auf entsprechende Fähigkeitsmerkmale einer Person geschlossen, ohne dass es dabei darauf ankommt, welche Items man betrachtet. Umgekehrt bedeutet dies, dass die Lösungswahrscheinlichkeit eines Items ausschließlich von zwei Faktoren abhängig ist: 1. von der Fähigkeit einer Person und 2. von der Schwierigkeit des Items. Dafür müssen aber weitere Modellvoraussetzungen erfüllt sein, die in Kapitel 4.7.1.2 näher beschrieben werden. Zunächst soll aber erklärt werden, wie das dichotome Rasch-Modell berechnet wird.

4.7.1.1 Das dichotome Rasch-Modell

Das **Rasch-Modell** ist eines der bekanntesten IRT-Modelle und versucht sicherzustellen, dass alle Testaufgaben auch wirklich die latente Variable in möglichst objektiver Weise messen (Bartolucci, Bacci & Gnaldi, 2016). Dabei wird außerdem berücksichtigt, dass nicht die Fähigkeit allein ausschlaggebend ist für das Testergebnis, sondern auch der Zufall (in Form einer Zufallsvariable U). Es spielt in den Item-Response-Theorien demnach auch die Wahrscheinlichkeit eine Rolle, mit der eine Person eine Aufgabe richtig beantwortet (die sog. Lösungswahrscheinlichkeit P). Der Grundsatz bleibt zwar bestehen, dass eine fähigere Person dabei eine Aufgabe eher lösen wird als eine weniger fähige Person, trotzdem kann auch die fähigere Person einen Flüchtigkeitsfehler begehen, der zur falschen Lösung führt. Zugleich hängt die Lösungswahrscheinlichkeit auch von der Schwierigkeit der Aufgabe ab. Bei leichten Aufgaben wird auch bei weniger fähigen Personen die Wahrscheinlichkeit hoch sein, dass die Aufgabe richtig gelöst wird. Somit beschreibt die Rasch-Modellgleichung die Wahrscheinlichkeit, dass eine Person mit einer bestimmten Fähigkeit (bezeichnet mit θ , θ) eine Aufgabe mit einer bestimmten Schwierigkeit (bezeichnet mit β , β) richtig beantwortet. Der Wert für die Wahrscheinlichkeit sollte dabei zwischen 0 und 1 liegen und außerdem die Funktion erfüllen, dass bei steigender Personen-Fähigkeit auch die Lösungswahrscheinlichkeit zunimmt. Diese Forderung ist für die logistische Funktion erfüllt (siehe Formel 1), weshalb sich die Modellgleichung des Rasch-Modells für $i = 1, \dots, n$ Personen und einem Test mit $j = 1, \dots, m$ Aufgaben wie folgt darstellt:

$$P(U_{ij} = 1 | \theta_i, \beta_j) = \frac{e^{\theta_i - \beta_j}}{1 + e^{\theta_i - \beta_j}} \quad (\text{Formel 1})$$

Man spricht hierbei vom Grundprinzip der **Maximum-Likelihood-Schätzung**, um die Item- und Personenparameter bestimmen zu können. Da das probabilistische Verfahren die Wahrscheinlichkeit mit einbezieht, dass ein berechneter Parameter auch falsch sein kann, so berechnet die Maximum-Likelihood-Schätzung mittels mathematischer Ableitung nach θ_i bei bekanntem Itemparameter β_j denjenigen Schätzwert der Personenfähigkeit (Maximum-Likelihood-Schätzer), bei welchem dieses Ereignis mit der größten Plausibilität (likelihood) tatsächlich auch eintritt. Da typischerweise die Itemparameter nicht vorab bekannt sind, müssen diese auf die gleiche Weise auf Basis der gesamten Datenmatrix geschätzt werden (Eid & Schmidt, 2014, S. 156ff.). Als vorteilhafte Variante hat sich hierbei die bedingte Maximum-Likelihood-Schätzung bewährt, weil sie im Vergleich zum unbedingten Verfahren nicht die Personen- und Itemparameter gemeinsam schätzt und somit mit steigender Personenstichprobe unter der Ungenauigkeit der Personenparameterschätzungen leidet, sondern die Itemparame-

ter unabhängig von der Schätzung der Personenparameter über die Anzahl der gelösten Aufgaben ermittelt: „Die Anzahl der gelösten Aufgaben [ist somit] eine *suffiziente Statistik* für die Personenfähigkeit. [...] Es kommt nur darauf an, *wie viele* Aufgaben eine Person gelöst hat, nicht aber *welche*“ (Eid & Schmidt, 2014, S. 162f.).

Den entscheidenden Teil der Formel stellt demnach die Differenz $\theta_i - \beta_j$ dar, welche sowohl im Zähler als auch im Nenner der Formel vorkommt. Die logistische Funktion (und somit die Lösungswahrscheinlichkeit) steigt, wenn der Exponent $(\theta_i - \beta_j)$ größer wird. Daraus kann man die gewünschte Abhängigkeit der Lösungswahrscheinlichkeit von der Aufgabenschwierigkeit gut nachvollziehen: Ist die Person fähiger als die Aufgabe schwer ist ($\theta_i > \beta_j$), ergibt sich ein positiver Wert und somit eine hohe Lösungswahrscheinlichkeit, ist die Aufgabe schwerer als die Person fähig ist ($\theta_i < \beta_j$), ergibt sich ein negativer Wert und somit eine kleine Lösungswahrscheinlichkeit. Die Lösungswahrscheinlichkeit steigt also mit der Überlegenheit der Person über die Aufgabe (Strobl, 2012, S. 5–9).

Bei der graphischen Darstellung der logistischen Funktion lässt sich eine Besonderheit erkennen, die sich durch den beschränkten Wertebereich ergibt: im mittleren Bereich liegt die größte Steigung vor, am negativen und positiven Ende flacht die Kurve ab (siehe Abb. 56). Man bezeichnet diese Darstellung als aufgabencharakteristische Kurve oder ICC (vom englischen Item Characteristic Curve), auf deren X-Achse die Personen-Fähigkeit θ_i und auf deren Y-Achse die Lösungswahrscheinlichkeit P aufgetragen wird. Im Beispiel der Abb. 56 hat demnach eine Person i mit Fähigkeit $\theta_i = 4$ eine wesentlich geringere Lösungswahrscheinlichkeit als eine Person mit Fähigkeit $\theta_i = 8$. Wenn die Lösungswahrscheinlichkeit 0,5 beträgt, ist die Person genauso fähig wie die Aufgabe schwer ist ($\theta_i = \beta_j$), d. h. die Wahrscheinlichkeit ist gleich groß, dass die Aufgabe von der Person gelöst werden kann oder auch nicht.

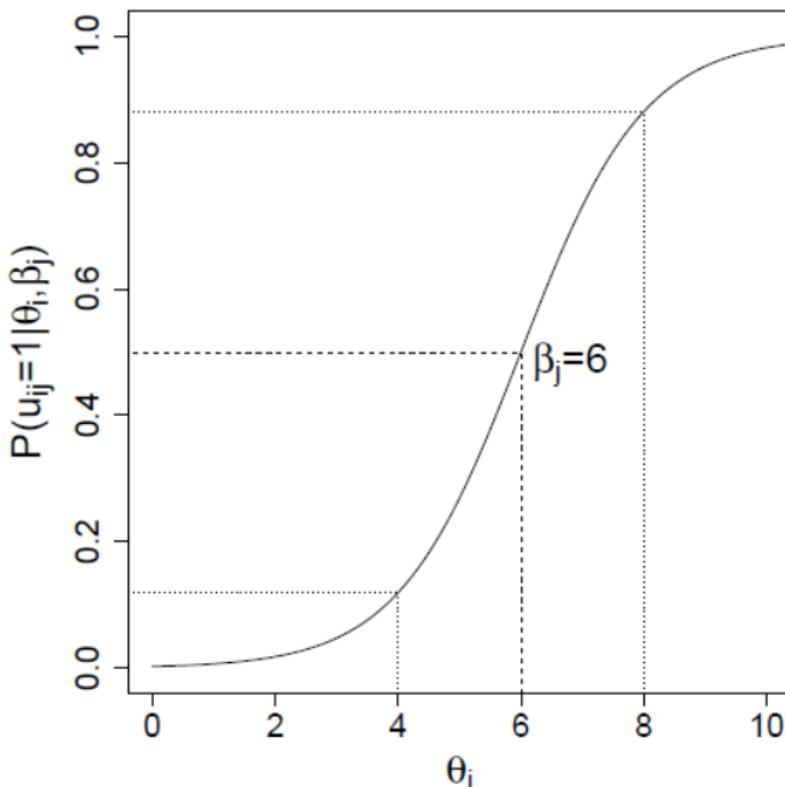
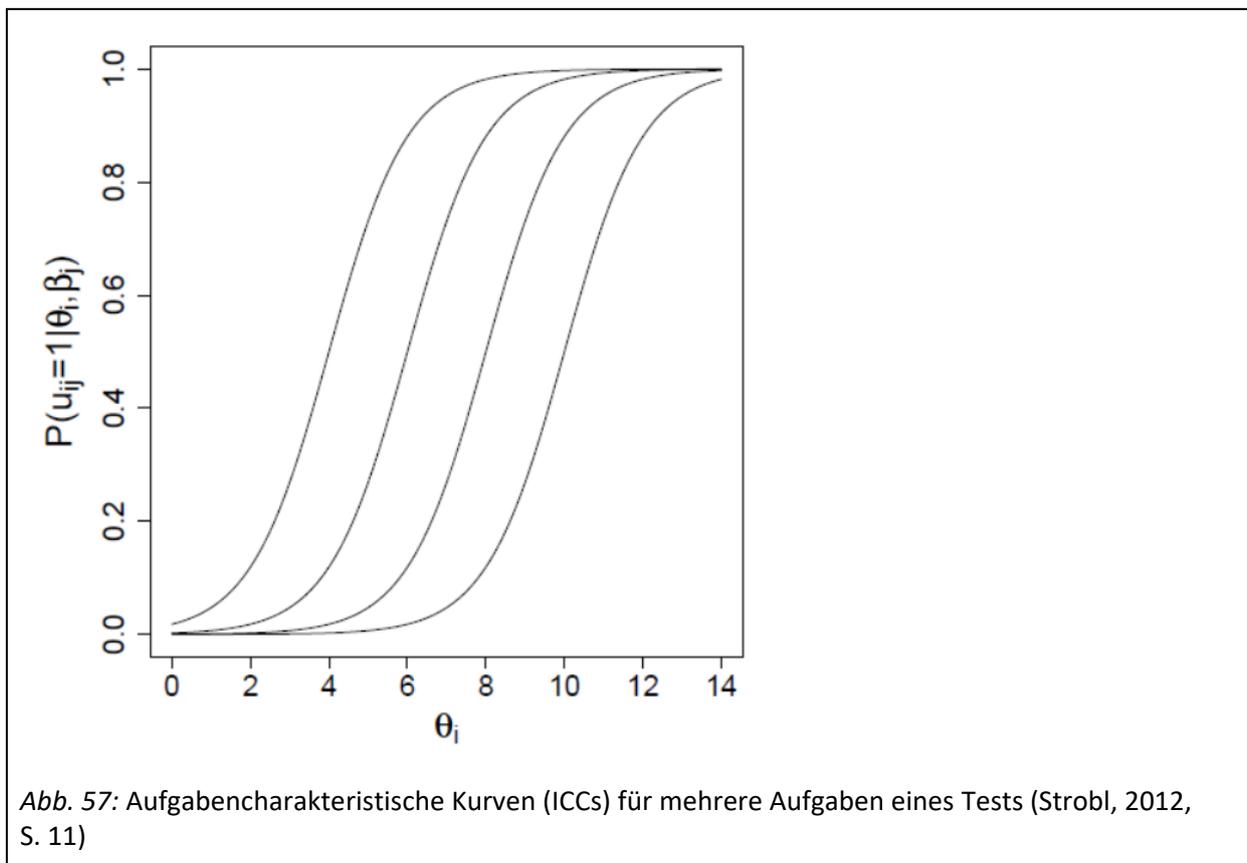


Abb. 56: Aufgabencharakteristische Kurve (ICC) für eine Aufgabe mit Schwierigkeit $\beta_j = 6$ (Strobl, 2012, S. 10)

Betrachtet man alle Aufgaben eines Tests, so fällt auf, dass beim Rasch-Modell alle ICCs parallel zueinander verlaufen und sich nicht kreuzen. Die Kurven können zwar je nach Aufgabenschwierigkeit nach links oder rechts verschoben sein, aber die Form ist immer dieselbe (siehe Abb. 57). Dies hängt damit zusammen, dass sich (nach der Modellgleichung) bei jeder Aufgabe nur der Parameter der Aufgabenschwierigkeit β_j ändert, wohingegen die Personen-Fähigkeiten der Probandengruppe gleichbleiben. Die über die einzelnen Aufgaben hinweg konstant bleibende Steigung der logistischen Funktion entspricht der Trennschärfe der jeweiligen Aufgaben, die im Idealfall möglichst groß sein soll, weil dann differenziert zwischen Personen unterschieden werden kann, selbst wenn ihre Fähigkeiten sehr ähnlich sind. Dies bedeutet gleichzeitig, dass eine Auswertung nach dem Rasch-Modell nur Aufgaben mit gleicher Trennschärfe zulässt, was eine sehr strenge Forderung in der Testkonstruktion darstellt.



4.7.1.2 Zentrale Annahmen und Eigenschaften des Rasch-Modells

Wenn ein Testverfahren den Anforderungen des Rasch-Modells entspricht, so ist dies ein starkes Argument für den empirischen Gehalt des Tests. Es gilt dabei vorab statistisch zu prüfen, ob die Eigenschaften für das Rasch-Modell erfüllt werden, damit man eben erwähnte Modellgleichung (Formel 1) auf die ermittelten Testdaten anwenden kann. Eine erste zentrale Voraussetzung des Rasch-Modells ist die **Eindimensionalität bzw. Homogenität** der Items. Items sind dann homogen, wenn sie alle das gleiche Konstrukt messen, d. h. alle Items beziehen sich auf eine einzige latente Variable (z. B. Fachwissen), ohne dass andere Einflussfaktoren vorliegen, welche die Lösungswahrscheinlichkeit des Items beeinflussen könnten (wie z. B. das Sprachverständnis beim Lesen einer Aufgabe). Homogenität zeigt

sich darin, dass diese Items untereinander eine positive Korrelation aufweisen, welche auf die Personenfähigkeit zurückzuführen ist. Personen mit höheren Fähigkeitswerten können demnach Items zuverlässiger lösen als Personen mit niedrigeren Fähigkeitswerten. Wenn im Test Items enthalten sind, die auf eine andere latente Variable zurückzuführen sind, so werden die Itemschwierigkeiten nicht mehr positiv miteinander korrelieren und die Homogenitätsannahme gilt als verletzt. Ob diese Korrelation zwischen den Items höher oder niedriger als erwartet ist, kann nicht anhand eines bestimmten Wertes beurteilt werden, sondern wird erst durch das Ergebnis der Modellprüfung sichtbar (Koller, Alexandrowicz & Hatzinger, 2012, S. 16f.).

Als Voraussetzung für Homogenität muss die **lokale Unabhängigkeit der Items** untereinander festgestellt werden. Dies bedeutet, dass alle Items für sich gelöst werden können, ohne dass das Lösen einer Aufgabe Voraussetzung für die Bewältigung nachfolgender Items sein darf. Jede Aufgabe kann also für sich allein bewältigt werden, ohne vorhergehende Aufgaben gelöst haben zu müssen. Die Lösungen dürfen nicht aufeinander aufbauen. Eine Verletzung der lokalen Unabhängigkeit liegt vor, wenn ein häufiges Auftreten der Antwortmuster {0 0} oder {1 1} zu beobachten ist, wie dies z. B. bei zwei inhaltlich ähnlichen oder aufeinander aufbauenden Items der Fall wäre, oder wenn Multidimensionalität zwischen zwei Items besteht (z. B. durch die Anwendung einer bestimmten Lösungsstrategie, die nichts mit der eigentlichen Fähigkeitsausprägung des latenten Konstruktes zu tun hat). Der Zusammenhang von Items untereinander darf sich ausschließlich durch die Fähigkeitsunterschiede der Testpersonen erklären lassen, oder umgekehrt, wenn die Fähigkeitsausprägung von Testpersonen hypothetisch konstant gehalten wird, so weisen die Items untereinander keinen Zusammenhang mehr auf (Rentzsch & Schütz, 2009, S. 153f.). Diese Forderung nach lokaler stochastischer Unabhängigkeit muss bei der Betrachtung einer einzelnen Person oder mehrerer Personen mit der gleichen Fähigkeit gelten, nicht aber für Personen mit unterschiedlichen Fähigkeitswerten. Ist dieser Grundsatz erfüllt, so darf man die Wahrscheinlichkeiten von Einzelereignissen multiplizieren, um deren gemeinsames Auftreten zu berechnen. Wenn diese Eigenschaft nicht erfüllt ist, zeigt sich dies in zu hohen Korrelationen zwischen den Items (Inter-Itemkorrelationen). Der Begriff lokal bezeichnet dabei die Tatsache, dass jeweils nur eine einzige Personenfähigkeit betrachtet wird, die stochastische Unabhängigkeit bezeichnet den Umstand, dass sich die zwei Zufallsereignisse Lösen und Nicht-Lösen von zwei Items nicht gegenseitig beeinflussen, d. h. die Wahrscheinlichkeit, dass das Item gelöst wird, ist unabhängig davon, ob das zweite Item gelöst wurde. Da lokal stochastisch abhängige Items keine zusätzlichen Informationen für die Bestimmung der Personenfähigkeit bringen, sollten sie aus dem Test eliminiert werden, wenn die Verletzung der lokalen Unabhängigkeit nicht durch Umordnung der Itemreihenfolge im Test aufgelöst werden kann (Koller et al., 2012, S. 18).

Ein weiterer wichtiger Grundsatz zur Anwendung des Rasch-Modells besteht in der Prüfung der **spezifischen Objektivität**. Es muss gewährleistet sein, dass Aussagen über die Fähigkeiten von zwei Personen nicht davon abhängen, anhand welcher Aufgaben sie verglichen werden. Dieser Grundsatz muss für alle Aufgaben erfüllt sein, unabhängig von der Wahl der Person. Es muss also immer gewährleistet sein, dass die Lösungswahrscheinlichkeit einer fähigeren Person bei jeder Testaufgabe (und somit unabhängig von der Aufgabenschwierigkeit) größer ist als die Lösungswahrscheinlichkeit einer weniger fähigen Person. Man spricht in diesem Fall auch von Stichprobenunabhängigkeit, weil man aus einem beliebigen Teil der Population Personen ziehen und damit die Personenfähigkeit schätzen kann, egal welche Personen (oder Items) dafür verwendet werden. Wenn sich eine Aufgabe für unterschiedliche Personen-Gruppen bei gleicher Fähigkeit als unterschiedlich schwer herausstellt, so spricht man von

Differential Item Functioning DIF (siehe Kapitel 4.7.1.5). Dies kann die Eliminierung der entsprechenden Aufgabe aus dem Test nach sich ziehen, weil Items dann die Voraussetzungen für die Rasch-Analyse nicht erfüllen (Koller et al., 2012, S. 19).

Technisch betrachtet kann die spezifische Objektivität nur durch die parallelen ICCs überprüft werden, die streng monoton steigen. Dies zeigt an, dass die Lösungswahrscheinlichkeit der Items mit steigender Personenfähigkeit durchgängig steigt und an keiner Stelle des Fähigkeitsbereichs konstant bleibt oder wieder absinkt. In Abb. 58 sieht man die Auswertung der Items des Fachwissenstests Teil A (Räumlich-zeitliches Wissen), welche die Voraussetzung der spezifischen Objektivität erfüllen. Strobl (2012, S. 80) stellt aber einschränkend fest: „Die Tatsache, dass die ICCs parallel dargestellt sind, bedeutet übrigens noch nicht, dass tatsächlich alle Aufgaben dieselbe Trennschärfe haben. Dies wird durch das Rasch-Modell zunächst nur angenommen.“ Daher muss in einem weiteren Schritt noch formal überprüft werden, ob das 2PL-Modell (Birnbaum-Modell), das unterschiedliche Trennschärfen zulässt, besser zu den Daten passt als das Rasch-Modell (siehe Kapitel 4.7.1.3).

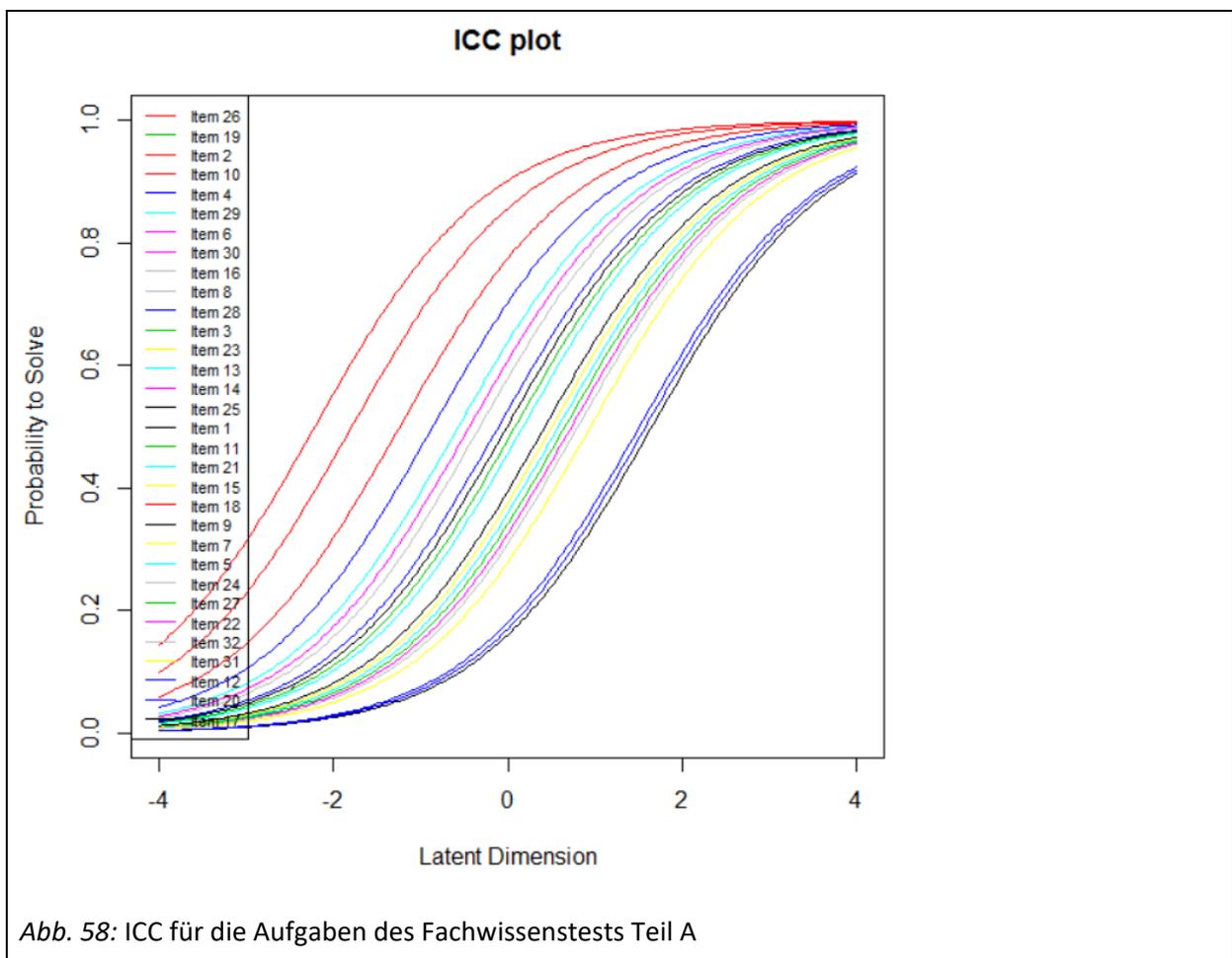


Abb. 58: ICC für die Aufgaben des Fachwissenstests Teil A

Ein zentraler Aspekt der spezifischen Objektivität der Personenfähigkeit ist die **Subgruppeninvarianz**, welche nachgewiesen werden kann, wenn sich die Schätzung der Personenfähigkeiten über unterschiedliche Subgruppen von Personen (z. B. Männer und Frauen) gleich verhält. Personen aus unterschiedlichen Gruppen, deren wahre Fähigkeit gleich hoch ist, erzielen durch die Beantwortung der Items in einem Test den gleichen Personenwert. Wenn die Items in unterschiedlichen Subgruppen

trotz gleicher Personenfähigkeit unterschiedlich beantwortet werden, so ‚funktionieren‘ die Items zwischen den Gruppen unterschiedlich und messen somit auch nicht dieselbe Fähigkeit. Dies kann darauf hindeuten, dass zum Beantworten der Items unterschiedliche Fähigkeitskomponenten erforderlich sind, die in den verschiedenen Personengruppen unterschiedlich ausgeprägt sind. Die Subgruppen können dabei anhand externer Variablen (z. B. Geschlecht) oder anhand interner Kriterien (z. B. Median des Personenparameters) geteilt werden, ehe die Itemparameterschätzungen der Gruppen miteinander verglichen werden (Koller et al., 2012, S. 20f.). Die Überprüfung der Subgruppeninvarianz anhand von externen Teilungskriterien wird in der Fachliteratur wie bereits erwähnt als die Überprüfung von Differential Item Functioning (DIF) bezeichnet (siehe Kapitel 4.7.1.5).

Wenn die eben beschriebenen Eigenschaften des Rasch-Modells für einen Datensatz zutreffen, dann ist gewährleistet, dass für jeden unbekannt Parameter eine **suffiziente Statistik** vorliegt, d. h. die gesamte relevante Information zur Schätzung der Personenparameter oder der Itemschwierigkeit ist in den entsprechenden Randsummen enthalten. Für die Personenfähigkeit enthält die Zeilen-Randsumme (d. h. die Addition aller bei einer Aufgabe erreichten Punkte) die gesamte Information über den Parameter θ_i . Die Information über den Parameter β_j ist in der Spalten-Randsumme enthalten, die ermittelt, wie viele Personen die jeweilige Aufgabe erfolgreich gelöst haben. Daraus lässt sich ableiten, „dass man nicht wissen muss, welche Aufgaben eine Person gelöst hat, um ihre Fähigkeit schätzen zu können, sondern nur wie viele Aufgaben sie insgesamt richtig gelöst hat“ (Strobl, 2012, S. 15).

4.7.1.3 Prüfung der Modellauswahl

Um das Modell mit der besten Modellanpassungsgüte für ein Testverfahren zu ermitteln, kann man auf informationstheoretische Maße zurückgreifen, bei denen die Likelihood des Modells mit der Anzahl der zu schätzenden Parameter verrechnet wird. Dabei soll ein gut passendes, aber zugleich sparsames Modell ausgewählt werden, d. h. je mehr Parameter frei geschätzt werden, desto weniger sparsam ist es, aber desto besser passt es, weil die Heterogenität in den Daten umso besser berücksichtigt werden kann. Die informationstheoretischen Maße Akaike Information Criterion (AIC) und Bayes Information Criterion (BIC) können bei der Auswahl eines Modells helfen, bei dem Modellanpassungsgüte und Sparsamkeit möglichst gut miteinander verknüpft sind.

Dabei stellt das zweiparametrische logistische Modell (Birnbaum-Modell, 2PL-Modell) eine Erweiterung des Rasch-Modells dar, indem der Diskriminationsparameter λ_j , welcher die Stärke des Items zur Unterscheidung zwischen Testpersonen mit unterschiedlicher Personenfähigkeit angibt, in die Formel des Rasch-Modells aufgenommen wird (Formel 2). Aufgrund der unterschiedlichen Diskriminationsparameter gibt es in diesem Modell schneidende ICC-Kurven, so dass die spezifische Objektivität in diesem Modell nicht gegeben ist (siehe Abb. 59). Die Verletzung der spezifischen Objektivitätsannahme bedingt, dass der Vergleich zwischen Personenparametern verschiedener Testpersonen von der Auswahl der Items abhängt, mit denen man die Fähigkeit gemessen hat. Im Umkehrschluss bedeutet dies auch, dass der Vergleich zwischen den Schwierigkeitsparametern der Items davon abhängt, welche Testpersonen man für die Berechnung der Itemschwierigkeit heranzieht. Dies hat dann zur Folge, dass im Gegensatz zum Rasch-Modell beim 2PL-Modell für Testpersonen mit derselben Anzahl richtig gelöster Aufgaben, aber einem unterschiedlichen Antwortenmuster unterschiedliche Personenparameter berechnet werden (Bartolucci et al., 2016). Sollte das 2PL-Modell nach Prüfung mittels der oben genannten informationstheoretischen Maße besser zu den Daten passen als das Rasch-Modell, so kann der Personenparameter nicht mit dem Rasch-Modell ermittelt werden.

$$P(U_{ij} = 1 | \theta_i, \beta_j) = \frac{e^{\lambda_j(\theta_i - \beta_j)}}{1 + e^{\lambda_j(\theta_i - \beta_j)}} \quad (\text{Formel 2})$$

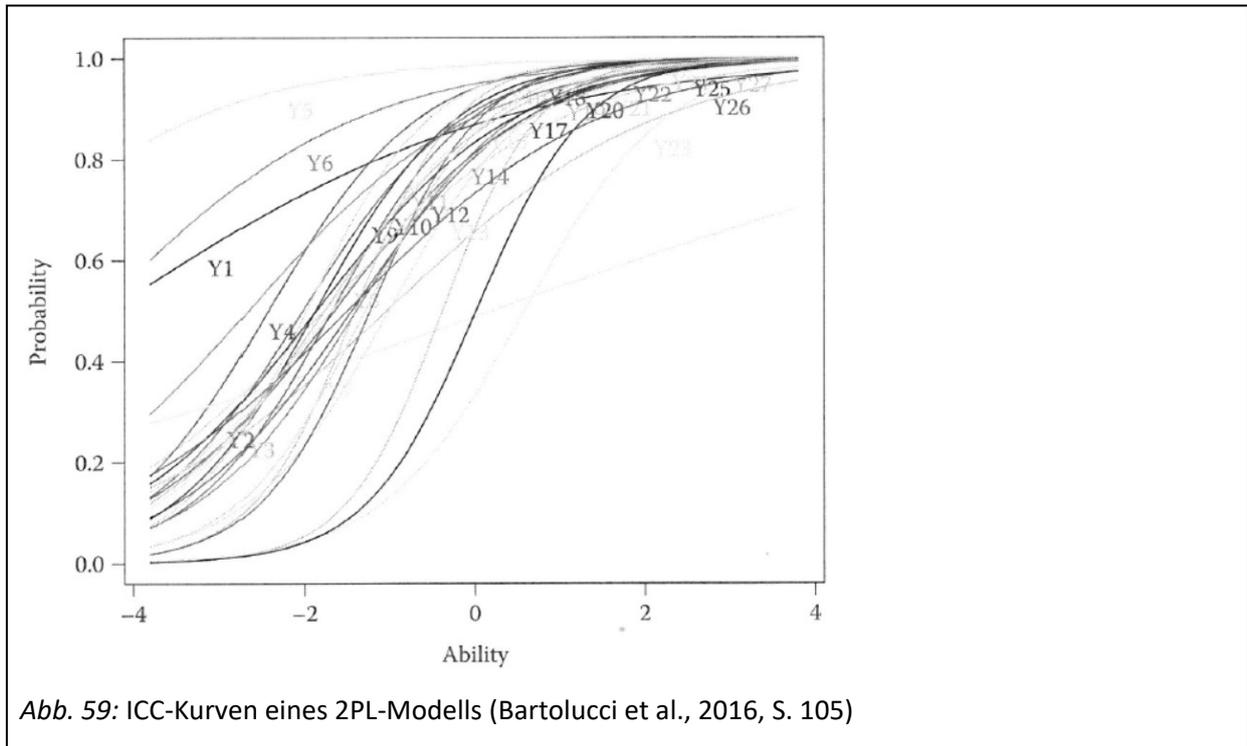


Abb. 59: ICC-Kurven eines 2PL-Modells (Bartolucci et al., 2016, S. 105)

Der Vergleich zwischen dem eindimensionalen Rasch-Modell (1PL-Modell) und dem zweiparametrischen logistischen Modell (Birnbbaum-Modell, 2PL-Modell) erfolgt mit dem R-package ltm (Rizopoulos, 2006). Dabei wird zunächst a priori im Einklang mit der Hypothese der Eindimensionalität lediglich eine latente Variable spezifiziert und die Log-Likelihood des auf die Datenmatrix angepassten Modells sowie die informationstheoretischen Indizes AIC und BIC sowie die geschätzten Itemparameter ausgegeben, die für den Modellvergleich interessant sind. Für die vorliegende Datenmatrix (data.mat) werden über den Befehl ltm(data.mat ~z1) zunächst die Maße für die Log-Likelihood sowie den AIC- und BIC-Index des 2PL-Modells berechnet. Anschließend erfolgt über den Befehl rasch(data=data.mat) die Berechnung dergleichen Maße für das eindimensionale Rasch-Modell (auch als 2PL-Modell mit identischen Ladungen bezeichnet). Mittels eines Likelihood-Ratio-Tests (über den R-Befehl anova) werden die Maße für beide Modelle dann in einer Tabelle (siehe Tab. 26) miteinander verglichen und geprüft, ob das 2PL-Modell mit frei schätzbaren Ladungen die Daten signifikant besser erklärt als das Modell mit identischen Ladungen (Rasch-Modell). Außerdem werden noch die χ^2 -verteilte Teststatistik, die Anzahl der Freiheitsgrade sowie die Überschreitungswahrscheinlichkeit berechnet. Es wird dann das Modell mit dem kleinsten Wert ausgewählt, weil dieses die Daten signifikant besser erklärt als das jeweils andere Modell (Bühner, 2011, S. 541). Wenn der Vergleich der AIC-Werte und der BIC-Werte unterschiedliche Ergebnisse liefert (wie dies auch in Tab. 26 der Fall ist), ist nach (Preinerstorfer & Formann, 2012) der BIC-Index dem AIC-Index überlegen, wenn es darum geht, das korrekte Modell aufzudecken, weil er nicht von der Stichprobengröße abhängig ist. Somit wird auch in dieser Arbeit der BIC-Index gegenüber dem AIC-Index favorisiert.

Tab. 26: Informationstheoretische Maße für die Teile E3 und E4

	AIC	BIC	log.Lik	LRT	Df	p.value
Rasch	917.63	944.64	-445.82			
2PL-Modell	905.32	955.19	-428.66	34.31	11	< 0.001

4.7.1.4 Statistische Prüfung der Modellgültigkeit

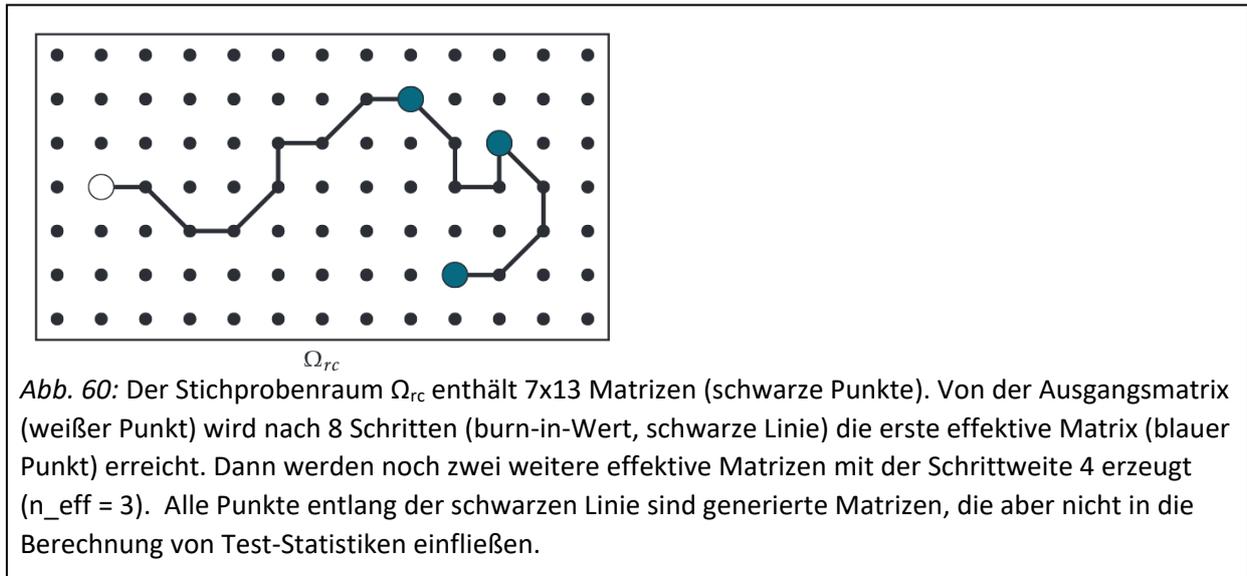
Es stellt sich nun die Frage, wie die in Kapitel 4.7.1.2 genannten Eigenschaften des Rasch-Modells statistisch überprüft werden können, um die entsprechenden Parameterschätzungen nach dem Rasch-Modell überhaupt durchführen zu können. Grundsätzlich lassen sich aus der Modellgleichung bestimmte Erwartungen hinsichtlich der Antwortmuster und -häufigkeiten treffen, die für den vorliegenden Datensatz zutreffen müssen, damit das Rasch-Modell verwendet werden darf. Das Prüfen dieser Annahmen geschieht über die Beantwortung spezifischer Fragen, beispielsweise ob für alle Aufgaben dieselbe Fähigkeit zur Lösung erforderlich ist oder bei einigen Aufgaben weitere Fähigkeiten eine Rolle spielen. Ebenso muss überprüft werden, ob die ICCs für alle Aufgaben tatsächlich parallel sind oder sich einzelne Items hinsichtlich ihrer Trennschärfe unterscheiden. Das Funktionieren der Aufgaben in allen Subgruppen kann ebenso ein Hinweis darauf sein, dass das Rasch-Modell die erhobenen Daten gut beschreibt. Als letztes kann über das Vorliegen lokaler stochastischer Unabhängigkeit die Modellgültigkeit dadurch nachgewiesen werden, indem gezeigt wird, dass der Erfolg oder Misserfolg bei einer Aufgabe keinen Einfluss auf die Lösungswahrscheinlichkeit anderer Aufgaben hat (Koller et al., 2012, S. 61).

Die standardmäßige Überprüfung der Modellgültigkeit erfolgt in der Regel mithilfe der gängigen parametrischen Berechnungen, die jedoch bei zu kleinen Stichproben einige statistische Schwierigkeiten nach sich ziehen können. Die Genauigkeit der Itemparameterschätzung beispielsweise steigt mit zunehmender Anzahl an Testwerten und kann bei zu geringer Probandenzahl zu Schätzungenauigkeiten (bis hin zur Unschätzbarkeit) des Parameters führen, welche dann wiederum in die Berechnung der Modellüberprüfung einfließen. Die parametrischen Modellgeltungstests beruhen außerdem auf asymptotischen Verteilungsannahmen (z. B. eine χ^2 -Verteilung beim Andersen-LRT), die nur bei hinreichend großen Stichproben gelten. Daher empfiehlt es sich, bei relativ geringem Stichprobenumfang, wie es in dieser Studie mit $N = 59$ der Fall ist, auf sogenannte **quasi-exakte Testverfahren** zurückzugreifen, welche nach Koller et al. (2012, S. 99) für die Überprüfung von Subgruppeninvarianz bereits ab $N = 50$ – 100 anwendbar sind (siehe Kapitel 4.7.1.4.1). Dabei ist es in der wissenschaftlichen Praxis durchaus üblich, nach der statistischen Berechnung mittels quasi-exakter Testverfahren (und der möglicherweise nötigen Eliminierung von Items) a posteriori Rasch-Modellgültigkeit für die verbleibenden Daten zu postulieren, was auch in dieser Studie zur Anwendung kommt. Trotz der eingeschränkten Anwendbarkeit der parametrischen Testverfahren wird deren Berechnung an geeigneten Stellen trotzdem durchgeführt, um die Ergebnisse der quasi-exakten Testverfahren zu überprüfen. So wird beispielsweise für den Beurteilungstest die Anpassungsgüte des Modells mit Hilfe einer Bootstrap-Goodness-of-Fit-Statistik mittels Pearson- χ^2 -Test über den R-Befehl `GoF.rasch` bewertet.

4.7.1.4.1 Nicht-parametrische Überprüfung mittels quasi-exakter Testverfahren

Die quasi-exakten Testverfahren können maximal zwei Gruppen von Personen untersuchen und gegenüberstellen, in welcher sich annähernd gleich viele Personen befinden müssen. Dies trifft beispielsweise in dieser Studie für die Trennung der Stichprobe nach ihrem Geschlecht (34 männliche, 25 weibliche Testpersonen) zu. Da die statistische Berechnung eines exakten Tests selbst bei relativ kleinen Datenmatrizen sehr schnell zu umfangreich wird, greift man auf Simulationsmethoden zurück, die eine Zufallsstichprobe von Datenmatrizen mit gleichen Randsummen erzeugen. Die Idee der quasi-exakten Tests beruht darauf, dass bei Gültigkeit des Rasch-Modells die Randsummen suffiziente Statistiken für die Personenfähigkeit und für die Itemschwierigkeit darstellen und somit alle möglichen binären Datenmatrizen zu gleichen Parameterschätzwerten führen. Daher müssen nicht alle möglichen Matrizen, welche die gleiche Randsumme haben wie die beobachtete Datenmatrix, zur Berechnung herangezogen werden, sondern es wird mit geeigneten Simulationsmethoden eine Zufallsstichprobe daraus gezogen. Die in dieser Studie verwendete Datensimulation orientiert sich an der Methode von Verhelst (2008), wie sie im R-package `RaschSampler` implementiert ist.

Die Schwierigkeit der Datensimulation besteht darin, dass bei der Simulation von Matrizen mit denselben Randsummen wie in der Ausgangsmatrix zufällig aus dem Raum aller möglichen Matrizen möglichst unähnliche Matrizen gezogen werden müssen, die zugleich mit gleicher Wahrscheinlichkeit erreicht werden können. Die Grundidee des Simulationsalgorithmus besteht somit darin, Einträge in der Matrix so zu vertauschen, dass die Personen- und Itemscores (Randsummen) gleichbleiben, und zugleich solche Matrizen auszuwählen, in welchen die Binärstrukturen möglichst stark voneinander abweichen. Die Unabhängigkeit und gleiche Wahrscheinlichkeit der Matrizen wird über die sog. Burn-in-Periode und die Schrittweite des Algorithmus gewährleistet (siehe Abb. 60). Der Burn-in-Wert legt fest, wie viele Matrizen ausgehend von der Ausgangsmatrix nur simuliert, dann jedoch nicht für die Berechnung des Modelltests genutzt werden. Nur die letzte entstandene Matrix wird verwendet und dient als Ausgangspunkt für die Erzeugung weiterer binärerer Matrizen, die sich auf diese Weise immer weiter von der Ausgangsmatrix entfernen. Somit ist die erste ausgewählte effektive Matrix der Ausgangsmatrix schon relativ unähnlich. Die Schrittweite (step) des Algorithmus legt fest, wie viele Vertauschungen von Matrixeinträge erfolgen müssen, bevor nach einer effektiven Matrix wieder eine neue Matrix für die Modellberechnung ausgewählt wird. Ein step-Wert von 32 bedeutet dann, dass 31 Matrizen verworfen und immer die 32. als nächste effektive Matrix ausgewählt wird. Als letztes wird noch festgelegt, anhand wie vieler simulierter Matrizen (n_{eff}) der p-Wert des Modelltests berechnet werden soll, d. h. wie viele effektive Matrizen mittels des eben erklärten Auswahlverfahrens simuliert und ausgewählt werden. Somit hat jede Matrix die gleiche Wahrscheinlichkeit, generiert zu werden, und die effektiv für den Modelltest verwendeten Matrizen entfernen sich immer weiter von der Ausgangsmatrix. Auf diese Weise kann vermieden werden, dass die für die Modellprüfung verwendeten Matrizen zu ähnlich zur Ausgangsmatrix sind und die Modellgültigkeit unberechtigter Weise nachgewiesen wird (Koller et al., 2012, S. 108–110). In einer weiteren Power-Analyse bestätigen Koller, Maier und Hatzinger (2015, o.S.): „The results showed that it is possible to investigate very small samples where a parametric approach is not possible, which is one of the most important advantages of quasi-exact tests.“



Die hierbei angewendeten Modelltests werden im R-package eRm über die Funktion NPtest() aufgerufen, wobei die Simulation der Matrizen gleichzeitig mit der Berechnung eines Modelltests (über den eingebundenen RaschSampler) gestartet wird. Für die Modelltestfunktion NPtest (method = "Tx") müssen vorab die eben erklärten Parametern n_{eff} (Anzahl effektiver Matrizen), burn_in (Abstandspareparameter zwischen den der Ausgangsmatrix und der ersten verwendeten effektiven Matrix), step (Schrittweite zwischen den ausgewählten effektiven Matrizen) und seed (Startwert für die Simulation) festgelegt werden. Das Problem fehlender Werte, die vor der Simulation aus der Ausgangsmatrix gelöscht oder durch geeignete Werte ersetzt werden müssten, stellt sich in dieser Studie nicht, da bei der Erhebung sorgfältig auf die Vollständigkeit der Daten geachtet wurde.

Grundsätzlich kann bei den quasi-exakten Testverfahren zwischen globalen Teststatistiken und solchen auf Itemebene unterschieden werden. Bei ersteren werden alle Items simultan überprüft, bei zweiten werden Itemgruppen mit mindestens zwei Items pro Subskala miteinander verglichen. Wenn die nachfolgenden Testverfahren positiv ausfallen, kann die Gültigkeit des Rasch-Modells für den ausgewählten Datensatz angenommen werden.

a) Überprüfung der Subgruppeninvarianz

Das erste angewendete quasi-exakte Testverfahren ist die **globale Teststatistik T_{10}** , welche überprüft, ob sich Items hinsichtlich ihrer Schwierigkeit zwischen Gruppen bedeutsam unterscheiden. Solche Tests zur Untersuchung von Subgruppeninvarianz erlauben Aussagen darüber, ob Items zwischen zwei oder mehreren Personengruppen unterschiedlich schwierig sind, indem der Datensatz beispielsweise anhand des externen Kriteriums Geschlecht in eine Referenz- und eine Zielgruppe geteilt wird. Zugleich lässt sich anhand eines internen Kriteriums (z. B. dem Personenwert-Median) die Annahme überprüfen, ob die ICCs monoton steigen oder ob die Trennschärfe des Items zu niedrig oder zu hoch ist. In R wird diese Aufteilung vorgegeben, indem nach Berechnung der Reihensumme (`score <- rowSums(datenmatrix)`) über den Befehl `splitvec <- ifelse (score <= median(score), 0, 1)` „Personen mit dem geringsten Personenscore bis einschließlich dem Personenscore am Median (\leq) der unteren Gruppe 0 und (else) alle Personen oberhalb des Personenscore-Medians der oberen Gruppe 1 zugeordnet werden“ (Koller et al., 2012, S. 54).

Rechnerisch zählt man die Personen in der Referenzgruppe, welche das eine Item i , jedoch nicht das Item j gelöst haben (Antwortmuster $\{1\ 0\}$, n_{ij}^{Ref}), und multipliziert diese Zahl mit der Anzahl von Personen in der Zielgruppe (Focal group), welche das Item j , jedoch nicht das Item i gelöst haben (Antwortmuster $\{0\ 1\}$, n_{ji}^{Foc}). Das Ergebnis subtrahiert man von den noch verbleibenden Möglichkeiten (Produkt aus den Anzahlen der Antwortmuster $\{0\ 1\}$ der Referenzgruppe und Antwortmuster $\{1\ 0\}$ der Zielgruppe, $n_{ji}^{Ref} \times n_{ij}^{Foc}$). Dieser Schritt wird für jede mögliche Itempaarkombination wiederholt. Die Absolut-Differenzen (ohne Beachtung des Vorzeichens der Differenz) jeder Itempaarkombination werden summiert und ergeben die Test-Statistik T_{10} (siehe Formel 3). Es geht nur die Anzahl von Matrizen in den Modelltest ein, in denen die absolute Differenzsumme gleich oder noch höher als in der Ausgangsmatrix ist. Die relative Häufigkeit dieser Anzahl von Matrizen wird im p -Wert ausgegeben, wobei $p > 0,05$ die Annahme der Modellgültigkeit für das entsprechende Splitkriterium zulässt. Die globale Teststatistik T_{10} überprüft somit ähnlich wie der parametrische Andersen-LR-Test den Datensatz auf Schwierigkeitsparameterdifferenzen zwischen zwei Gruppen, wobei das Verhältnis der untersuchten Antwortmuster zwischen den zwei Gruppen ungefähr gleich ist ($n_{ij}^{Ref} / n_{ji}^{Ref} \approx n_{ij}^{Foc} / n_{ji}^{Foc}$), wenn eine Itemkombination keine Verletzung der Subgruppeninvarianzannahme zeigt.

$$T_{10}(A) = \sum_{ij} \left| n_{ij}^{Ref} n_{ji}^{Foc} - n_{ji}^{Ref} n_{ij}^{Foc} \right| \quad (\text{Formel 3})$$

Es bleibt zu beachten, dass ein „nicht signifikantes Ergebnis noch nicht bedeutet, dass keine Verletzungen auf Itemebene vorhanden sind, denn durch das Aufsummieren über alle Items können wichtige Informationen bei der globalen Statistik verdeckt werden“ (Koller et al., 2012, S. 138). Dies wäre dann der Fall, wenn zwei Items die Annahme der Subgruppeninvarianz verletzen, jedoch die übrigen Items in den beiden Gruppen so gut funktionieren, dass der Differenzbetrag der beiden auffälligen Items zu wenig Auswirkung auf die gesamte Summe an Differenzbeträgen zeigt und der globale p -Wert nicht-signifikant bleibt.

Um nun genauer herauszufinden, welche Items der Modellannahme der Subgruppeninvarianz widersprechen, kommt das **Testverfahren T_4** auf Itemebene zur Anwendung. Dazu wird der Datensatz der Ausgangsmatrix anhand eines Teilungskriteriums wieder geteilt und nur für ein einzelnes Item gezählt, wie oft es in der Zielgruppe gelöst wird. Dieses Ergebnis stellt bereits die Test-Statistik T_4 der Ausgangsmatrix dar. Das gleiche Auszählen wird für alle simulierten Datenmatrizen durchgeführt. Um die Annahme zu prüfen, dass das Item zu leicht für die Zielgruppe ist, werden nur diejenigen Matrizen in die Modellberechnung einbezogen, die im Vergleich zur Test-Statistik der Ausgangsmatrix eine gleiche oder größere Lösungsanzahl aufweisen. Es wird also herausgefunden, ob ein Item in einer bestimmten Gruppe leichter oder schwieriger ist als unter dem Rasch-Modell erwartet, d. h. ob es zu viele oder zu wenige positive Antworten innerhalb von Personengruppen gibt. Wenn der Datensatz zur Berechnung nach dem Geschlecht in zwei Personengruppen unterteilt wird, bedeutet ein signifikantes Ergebnis, dass das Item für die untersuchte Gruppe (z. B. Frauen) signifikant leichter oder schwieriger ist als für die Referenzgruppe (z. B. Männer). Wendet man ein internes Teilungskriterium (z. B. den Personenscore-Median³⁸) zur Gruppentrennung an, so zeigt der p -Wert eine zu hohe oder zu niedrige Trennschärfe eines Items an, d. h. das Item ist in der unteren Fähigkeitsgruppe zu leicht oder in der oberen Gruppe zu schwierig. Es könnte jedoch auch eine erhöhte Ratewahrscheinlichkeit in der unteren oder eine zu hohe Irrtumswahrscheinlichkeit in der oberen Fähigkeitsgruppe dafür verantwortlich

³⁸ Der Median stellt den Personenscore dar, über dem und unter dem sich jeweils 50 Prozent der Testpersonen befinden.

sein, denn diese Arten von Modellverletzung sind über diese Teststatistik nicht unterscheidbar. Ein p-Wert von $p < 0,05$ zeigt eine Verletzung der Annahme der Subgruppeninvarianz, was allerdings nicht bedeutet, dass die untersuchte Gruppe dem Item leichter oder schwerer zustimmt als eine Referenzgruppe. Die T_4 -Statistik überprüft somit die Annahme der Subgruppeninvarianz und der Monotonizität der ICCs auf Itemebene. Um zu vermeiden, dass sich die berechneten Werte mehrerer Items in einer Itemgruppe gegenseitig kompensieren, wird T_4 für jedes zu überprüfende Item einzeln berechnet.

b) Überprüfung der lokalen stochastischen Unabhängigkeit und Homogenität

Die Überprüfung der lokalen stochastischen Unabhängigkeit zwischen Items (global test-local dependence) und deren Homogenität kann in einem ersten Schritt über die globale **Teststatistik T_{11}** (Ponocny, 2001) erfolgen. Inhaltlich werden im gesamten Datensatz unpassende Inter-Itemkorrelationen zwischen allen möglich Itempaaren identifiziert (d. h. es wird überprüft, ob zu niedrige oder zu hohe Inter-Itemkorrelationen vorhanden sind). Rechnerisch wird für die Ausgangsmatrix sowie für alle simulierten Matrizen für jedes mögliche Itempaar eine Korrelation berechnet. Diese Korrelationswerte aller simulierten Matrizen werden für jedes Itempaar miteinander addiert und (durch Division mit der Anzahl der simulierten Matrizen) die mittlere Inter-Itemkorrelation der entsprechenden Itempaare berechnet. Die absolute Differenz zwischen der empirisch vorliegenden Korrelation eines Itempaares der Ausgangsmatrix und dessen berechneter mittlerer Korrelation der simulierten Matrizen stellt die Abweichung zu der unter dem Rasch-Modell erwarteten Korrelation dar. Bei größer werdenden Abweichungen zur mittleren Korrelation steigt somit auch die Summe der Differenzen. „Je größer diese summierte absolute Abweichung ausfällt, desto stärker fällt die Verletzung der lokalen Unabhängigkeit aus“ (Bühner, 2011, S. 533). Die Differenzwerte aller möglichen Itempaare werden in der Teststatistik T_{11} gespeichert. Für den Modelltest wird ausgezählt, wie oft in den simulierten Matrizen die gleiche oder eine noch größere Abweichungssumme als in der Ausgangsmatrix gefunden wurde. Dieses Ergebnis wird dann durch die Anzahl der simulierten Matrizen dividiert, woraus der p-Wert für den Modelltest T_{11} resultiert. Wenn es nur wenige Differenzwerte der simulierten Matrizen gibt, die eine gleiche oder noch größere Abweichung als die Ausgangsmatrix besitzen, so wird das Ergebnis des Modelltests T_{11} höchstwahrscheinlich signifikant (d. h. der p-Wert kleiner 0,05) sein und die Items als lokal stochastisch abhängig zueinander beurteilt werden. Die Teststatistik kann aber nicht unterscheiden, ob eine Verletzung der lokal stochastischen Unabhängigkeit zwischen Items oder eine mögliche Heterogenität von Items vorliegt (Koller et al., 2012, S. 106ff.).

Mit der **Teststatistik T_{md}** (md bedeutet multidimensional scales) wird in einem zweiten Schritt ähnlich dem parametrischen Martin-Löf-Test global überprüft, ob zwei Subskalen eines Tests eindimensional messen, indem ihre jeweiligen Personenwerte positiv korreliert sind. Als einheitliches Splitkriterium für diesen Modelltest wird die Aufgabenschwierigkeit verwendet, indem die Items in zwei möglichst gleichgroße Gruppen anhand ihrer absoluten Lösungshäufigkeit geteilt werden (Itemscore-Median). Zunächst wird wiederum der Korrelationskoeffizient zwischen den beiden Subskalen der beobachteten Ausgangsmatrix bestimmt und dann mit den entsprechenden Korrelationskoeffizienten der simulierten Matrizen verglichen. Laut Modelltest werden nun wieder nur diejenigen Matrizen zur Berechnung herangezogen, in denen die Korrelation der beiden Personenscores gleich oder kleiner als in der beobachteten Ausgangsmatrix ist. Aus der relativen Häufigkeit dieser Anzahl ergibt sich dann der p-Wert, der mit einem Signifikanzniveau von $\alpha = 0,05$ verglichen wird. Wenn keine Verletzung ($p > 0,05$ ist nicht-signifikant) vorliegt, dann weisen weniger fähige Personen in beiden Subskalen eher geringe Personenwerte auf, welche mit steigender Personenfähigkeit in beiden Gruppen größer werden, d. h. die

Lösungswahrscheinlichkeit der Items steigt mit höher werdender Personenfähigkeit (Koller et al., 2012, S. 118ff.). Somit kann nicht nur die Eindimensionalität der Items, sondern auch die Monotonizitätsannahme der ICCs nachgewiesen werden.

Auf Itemebene ermöglicht die **Teststatistik T_1** eine Überprüfung der Annahme der lokalen stochastischen Unabhängigkeit, indem zunächst für jedes mögliche Itempaar die Auftretenshäufigkeit der Antwortmuster {0 0} (beide Items werden mit 0 beantwortet) und {1 1} (beide Items werden mit 1 beantwortet) in der beobachteten Ausgangsmatrix gezählt und summiert werden. Diese Anzahl stellt bereits die Test-Statistik T_1 dar und wird mit der Auftretenshäufigkeit der Antwortmuster {0 0} und {1 1} in allen simulierten Matrizen verglichen. Für die Berechnung wird nur die Anzahl von Matrizen verwendet, in denen die relative Auftretenshäufigkeit gleicher Antwortmuster gleich oder noch höher als in der Ausgangsmatrix ist. Diese relative Auftretenshäufigkeit gibt uns den p-Wert aus, der mit einem Signifikanzniveau von $\alpha = 0,05$ verglichen wird. Wenn für ein Itempaar der p-Wert kleiner 0,05 ist, so spricht dies für ihre signifikante Abhängigkeit zueinander und somit für eine Modellabweichung.

Die **Testverfahren T_{1m} und T_{1l}** stellen Modifikationen von T_1 dar. T_{1m} ($m = \text{multidimensional}$) überprüft, ob zu wenige gleiche Antwortmuster ({0 0} oder {1 1}) zwischen den möglichen Itemkombinationen auftreten. Sofern die Items homogen sind, sollte eine gewisse Abhängigkeit zwischen den Items vorhanden sein, d. h. sie korrelieren leicht positiv miteinander. Ist dies nicht der Fall, so werden wenige bis gar keine gleichen Antwortmuster vorhanden sein. T_{1l} ($l = \text{learning}$) überprüft, ob zu viele gleiche Antwortmuster {1 1} auftreten, ohne dass sich das Antwortmuster {0 0} erhöht. Dies könnte zu der Vermutung führen, dass durch das Lösen eines Items ein darauffolgendes Item leichter gelöst werden könnte (Koller et al., 2012, S. 127f.).

c) Überprüfung der Trennschärfe

Die Überprüfung der Trennschärfe kann bei geringer Stichprobengröße mit der **Teststatistik T_{pbis}** über die Software R erfolgen, welche die direkte Untersuchung der Trennschärfe von einzelnen Items im Vergleich zu der restlichen Skala erlaubt. Das Verfahren beruht auf der punkt-biserialen Korrelation ($pbis$), die in der klassischen Testtheorie für die Überprüfung der Trennschärfe bei dichotomen Items verwendet wird. Dazu wird der Datensatz für das zu untersuchende Item entsprechend dem Antwortverhalten in eine Gruppe 0, welche das Item nicht gelöst hat, und eine Gruppe 1, welche das Item gelöst hat, unterteilt. Mit dem Datensatz der restlichen Items wird dann der Personenparameter jeder Testperson berechnet und die Personenscores entsprechend der Gruppenzugehörigkeit addiert. Zum Schluss wird die Differenz der summierten Personenscores, welche mit der Personenanzahl der jeweils anderen Gruppe gewichtet werden, gebildet (siehe Formel 4).

$$T_{pbis}(A) = n_1 \sum r_0 - n_0 \sum r_1 \quad (\text{Formel 4})$$

Wenn die zwei Personengruppen gleich groß sind und sich in der Gruppe der Nicht-Löser eher Personen mit niedrigen Personenparametern und in der Gruppe der Löser eher Personen mit hohen Personenscores befinden, so wird T_{pbis} einen negativen Wert annehmen. Wenn der Wert positiv wird, so deutet das darauf hin, dass eher Personen mit hoher Fähigkeit das Item falsch beantwortet haben. T_{pbis} wird für die beobachtete Ausgangsmatrix sowie für alle simulierten Matrizen berechnet und anschließend wieder ausgezählt, wie viele simulierte Teststatistiken einen gleichen oder kleineren Differenzwert als in der beobachteten Matrix zeigen. Ist dies oft der Fall, dann ist der beobachtete Wert klein

genug und die Trennschärfe kann als gut angenommen werden (Koller et al., 2012, S. 150ff.). Der p -Wert berechnet sich aus der relativen Häufigkeit dieser Anzahl von Matrizen und zeigt bei einem Wert von $p > 0,05$, dass das Item trennscharf im Vergleich zu den anderen Items ist. Eine zu niedrige Trennschärfe kann aber ebenso bedeuten, dass das Item eine andere Fähigkeit anspricht und damit keine Homogenität der Items vorliegt, wodurch inhaltlich mit der Teststatistik auch das multidimensionale Funktionieren der Items geprüft werden kann.

d) Alpha-Korrektur und multiples Testen

Bei der in dieser Studie angewendeten Modellprüfung geht es darum zu beweisen, dass der vorliegende Datensatz für das Rasch-Modell zutrifft. Im Mittelpunkt der Datenanalyse steht das Zutreffen der Nullhypothese, dass keine Modellverletzung vorliegt. Ein Fehler erster Art liegt vor, wenn eine richtige Nullhypothese abgelehnt und eine ungültige Alternativhypothese angenommen wird. Das Risiko dieser falschen Entscheidung (Irrtumswahrscheinlichkeit) wird über das Signifikanzniveau α festgelegt und beträgt in der Regel 0,05 bzw. 5 % (signifikantes Ergebnis) oder 0,01 bzw. 1 % (hoch signifikantes Ergebnis). Einen Fehler zweiter Art begeht man, wenn man eine ungültige Nullhypothese beibehält, obwohl in Wirklichkeit eine Alternativhypothese richtig gewesen wäre. Zwischen diesen beiden Fehlern besteht ein indirekter Zusammenhang, denn je größer man den Fehler erster Art festlegt, umso kleiner wird der Fehler zweiter Art. Bei einer Modellprüfung mittels Itemanalyse geht es darum, ein Set von Items zu finden, das dem Rasch-Modell entspricht und die Nullhypothese (das Rasch-Modell gilt) bestätigt. Da die Nullhypothese nicht aufgrund von Zufälligkeiten leicht verworfen werden soll, muss die Wahrscheinlichkeit, eine ungültige Nullhypothese beizubehalten, gesenkt werden. Aus diesem Grund empfehlen Koller et al. (2012), für die Modellprüfung den Fehler erster Art auf $\alpha = 0,1$ (10 %) festzulegen.

Da beim Rasch-Modell mehrere Hypothesen mit verschiedenen Tests an einem Datensatz überprüft werden (z. B. die Annahme der Subgruppeninvarianz, der Homogenität der Items und der lokalen stochastischen Unabhängigkeit), spricht man von multiplem Testen. Das α -Risiko für eine Fehlentscheidung kumuliert sich, wenn die Nullhypothese mehrere Tests abgelehnt oder angenommen werden können. Die Wahrscheinlichkeit für einen Fehler erster Art ist also nicht mehr α , sondern größer, denn das Ereignis, dass die richtige Nullhypothese fälschlicherweise abgelehnt wird, kann bei jedem Einzeltest eintreten. Man spricht von α -Inflation oder α -Aufblähung, weil in Wirklichkeit ein viel höheres Risiko besteht, eine Fehlentscheidung zu treffen, als dies das festgelegte Signifikanzniveau vorgibt. Für die Modellprüfung bedeutet dies, dass man viel zu leicht annimmt, dass bestimmte Modellannahmen verletzt sind, obwohl sie in Wirklichkeit erfüllt sind. Daher muss das α für einen Einzeltest so korrigiert werden, dass das globale α (d. h. das α für alle Tests gemeinsam betrachtet) nicht größer als das entsprechend vorgegebene Signifikanzniveau ist. Aus diesem Grund wird bei den in dieser Studie verwendeten Modellprüfungen die Bonferroni-Korrektur angewandt, bei der ein korrigiertes α^* berechnet wird, indem das ursprüngliche α durch die Anzahl der durchgeführten Tests q dividiert wird. Diese Alpha-Korrektur wird allerdings nicht auf die gesamte Modellprüfung angewendet, sondern nur innerhalb der einzelnen Teststatistiken der Modelltests durchgeführt, weil sonst die Gefahr besteht, dass das Rasch-Modell zu leicht fälschlicherweise als gültig angenommen wird. Die Alpha-Korrektur wird für die einzelnen quasi-exakten Testverfahren anhand der Empfehlungen von Koller et al. (2012) durchgeführt, die in Tab. 27 dargestellt sind.

Tab. 27: Alpha-Korrektur für die quasi-exakten Tests nach Koller et al. (2012)

Subgruppeninvarianz	
T_{10}	α/q
T_4 explorativ	α/q
T_4	$\alpha/(q \times \text{Anzahl durchgeführter Tests})$
LSA und/oder Multidimensionalität	
T_{11}	keine Korrektur
$T_1, T_{1m}, T_{1\ell}$	$\alpha/\text{Anzahl Itemgruppen}$
T_2, T_{2m}	$\alpha/\text{Anzahl durchgeführter Tests}$
Mehrdimensionale Subskalen	
T_{md}	α/q
Diskrimination	
T_{pbis}	α/k

Anmerkung: α ... Alpha; q ... Anzahl Teilungskriterien; k ... Anzahl der Items im Test; LSA ... Lokale stochastische Abhängigkeit.

4.7.1.4.2 Parametrische Berechnungen

Die parametrische Überprüfung der Modellgültigkeit stellt zwar die gängigste Möglichkeit dar, um die Anwendung des Rasch-Modells auf einen vorliegenden Datensatz zu rechtfertigen, wird jedoch im Rahmen dieser Studie nur zur Absicherung der quasi-exakten Testverfahren angewendet. Damit beispielsweise bei einem etablierten Verfahren wie dem **Pearson- χ^2 -Test** die Prüfgröße nämlich einer χ^2 -Verteilung folgt, muss die Stichprobe entsprechend groß sein, um sicherzustellen, dass die erwarteten Häufigkeiten pro Antwortpattern mindestens 1 betragen. Um dies für kleine Stichproben sicherzustellen, kann zwar auf parametrische Bootstrap-Verfahren zurückgegriffen werden, in welchen aus bereits geschätzten Parametern anhand eines Simulationsverfahrens wiederholt Stichproben gleicher Größe gezogen und der χ^2 -Wert bestimmt wird. Auch wenn jedoch grundsätzlich als einzige Voraussetzung des Bootstrap-Verfahrens erfüllt sein muss, dass die Stichprobe repräsentativ für die Population ist, haben Langeheine, Pannekoek und de Po (1996) herausgefunden, dass erst mit minimal 300 Stichproben zufriedenstellende Ergebnisse bei parametrischen Bootstrap-Verfahren erzielt werden können. Bühner (2011, S. 537) lehnt den Pearson- χ^2 -Test sogar grundsätzlich ab, weil die Voraussetzungen für diesen Test praktisch nie gegeben sind, denn dazu müsste jedes Antwortmuster mindestens einmal beobachtet werden. Das Bootstrap-Verfahren habe sich vor allem bei dichotomen Items ebenso als wenig teststark erwiesen, um Modellverletzungen in Form von sich überschneidenden ICCs zu erkennen. Daher kann lediglich bei signifikanten Werten mit großer Wahrscheinlichkeit angenommen werden, dass keine Raschkonformität vorliegt, wohingegen ein nicht-signifikantes Ergebnis noch kein zuverlässiger Nachweis für die Annahme des Raschmodells darstellt.

Mit dem Andersen-Likelihood-Ratio-Test (LRT) wird die globale Gültigkeit des Raschmodells überprüft, indem alle Items simultan in die Berechnung der Likelihood (Produkt der Wahrscheinlichkeiten für die beobachteten Daten bei gegebenen Parametern) miteinfließen. Dabei wird zuerst die Likelihood für den gesamten Datensatz, dann die Likelihoods für zwei oder mehrere getrennte Gruppen berechnet. In dieser Studie wird die Stichprobe beispielsweise nach den Splitkriterien Median oder Geschlecht in zwei Gruppen unterteilt. Der Quotient aus der Gesamtl likelihood bei getrennter Schätzung und jener

aus der gemeinsamen Schätzung besagt, um wie viel Mal die getrennte Schätzung besser ist. Je größer der Quotientenwert ist, umso eher werden sich Itemschwierigkeiten in den einzelnen Gruppen systematisch unterscheiden. Wenn sich die Likelihoods aus der Gesamtgruppe und den Teilgruppen nicht unterscheiden, dann darf angenommen werden, dass auch die Parameterschätzungen für die Gesamtgruppe und die Teilgruppen gleich sind. Ein signifikantes Ergebnis ($p < 0,05$) deutet darauf hin, dass die Nullhypothese (H_0 : Gültigkeit des Rasch Modells wird angenommen) nicht gilt.

Je nach Teilungskriterium kann der LRT unterschiedliche Arten von Modellannahmen überprüfen. Wenn die Personengruppen anhand des Personenscore-Medians geteilt wird, so kann eine Einschätzung darüber getroffen werden, ob sich die Itemschwierigkeit zwischen der unteren und der oberen Fähigkeitsgruppe unterscheidet. Ein signifikantes Ergebnis des LRT bedeutet, dass Items existieren, die in einer der beiden Gruppen viel schwieriger oder leichter sind als in der Referenzgruppe. Wenn das Ergebnis nicht-signifikant ist, kann man davon ausgehen, dass die Lösungswahrscheinlichkeit der Items mit steigender Personenfähigkeit ebenso steigen, was wiederum für den einheitlichen Anstieg der ICCs spricht. Wählt man hingegen ein externes Splitkriterium wie das Geschlecht, so spricht ein signifikanter LRT-Wert für das Vorliegen von DIF in Form von Subgruppeninvarianz (Koller et al., 2012, S. 67–76).

Das Rasch-Modell beinhaltet die Tatsache, dass die Itemparameter in allen Subpopulationen (z. B. bei Männern und Frauen) denselben Wert aufweisen müssen. Der **graphische Modelltest** (erstellt mit dem lrt-package über den Befehl plotGOF) überträgt die in zwei verschiedenen Subgruppen geschätzten Itemparameter auf die beiden Achsen eines Koordinatensystems. Bei Modellgültigkeit dürfen sich die geschätzten Werte nur aufgrund des Schätzfehlers unterscheiden und liegen im Idealfall auf einer Winkelhalbierenden. Ob sich die Parameterschätzungen bedeutsam unterscheiden, kann man überprüfen, wenn die Winkelhalbierende bei einzelnen Items außerhalb der 95 %-Konfidenzellipse oder das Item außerhalb des 95 %-Konfidenzbandes liegt, wie dies in Abb. 61 bei Item beta D6.3 des Beurteilungstests der Fall ist (Eid & Schmidt, 2014, S. 183–185).

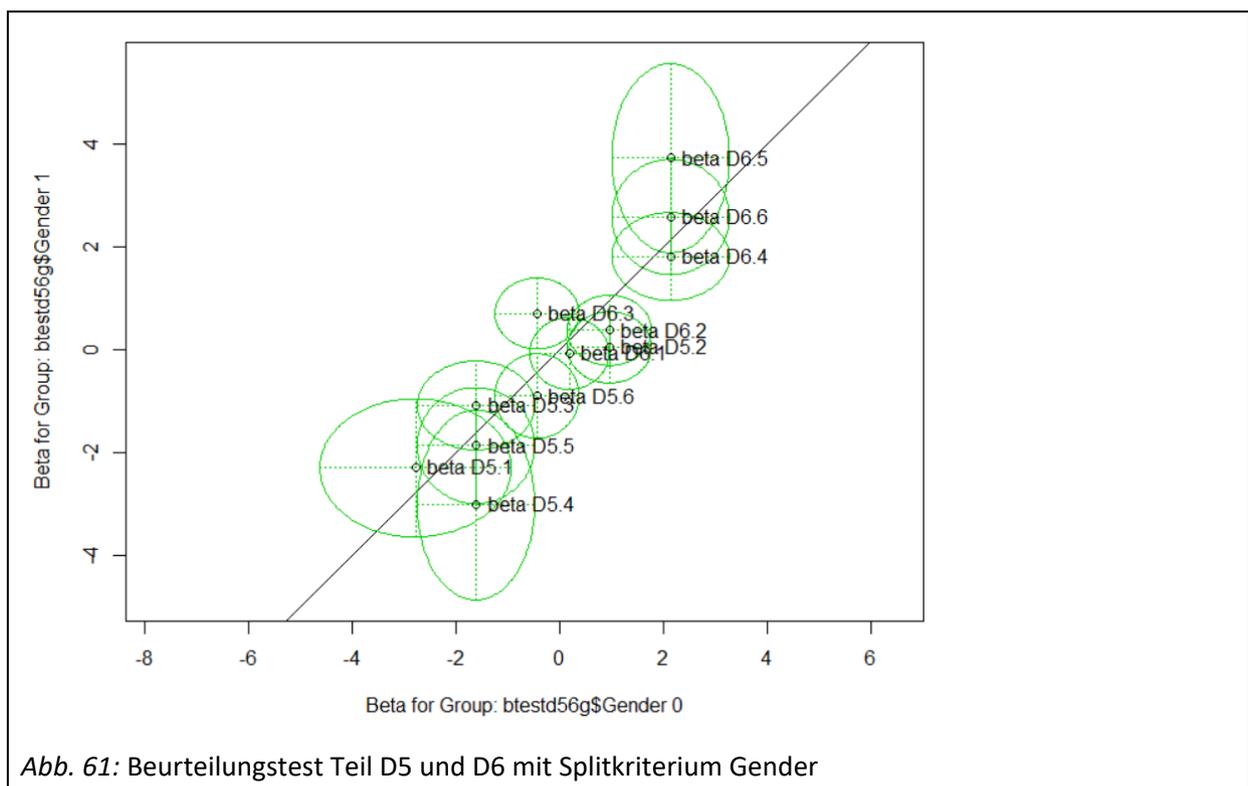


Abb. 61: Beurteilungstest Teil D5 und D6 mit Splitkriterium Gender

Eine ähnliche Subgruppenuntersuchung liefert der **Martin-Löf-Test** für die Annahme der Homogenität von Itemsuskalen, indem die Likelihoods aus dem Gesamttest mit den beiden Subtests verglichen werden:

Ein signifikantes Ergebnis dieses Tests zeigt an, dass die Personenscores nicht eindimensional sind und daher die Items der beiden Testhälften unterschiedliches messen. [...] Erfassen die Itemsuskalen die gleiche Dimension, dann steigt bei Konstanthaltung der Personenreihenfolge und Anstieg der Fähigkeit die Anzahl der gelösten Items in allen untersuchten Subgruppen. (Koller et al., 2012, S. 94f.)

Wendet man für den Martin-Löf-Test ein internes Teilungskriterium wie den Median an, so ermöglicht dies die Entdeckung einer möglichen Multidimensionalität zwischen Itemsuskalen oder auch den unterschiedlichen Anstieg der ICCs auf globaler Ebene. Dies wiederum könnte darauf hindeuten, dass mit steigender Fähigkeit in der einen Subkala die Lösungswahrscheinlichkeit für Items aus der anderen Subkala nicht steigt. Der Test benötigt allerdings auch relativ große Stichproben (mit mehreren hundert Testpersonen), um zu zuverlässigen Ergebnissen zu gelangen, weil er auf den relativen Häufigkeiten der einzelnen Personenscore-Kombinationen beruht (Koller et al., 2012, S. 90–97).

Eine andere Möglichkeit, die Subgruppeninvarianz der Itemparameter grafisch darzustellen, ist der Plot der Konfidenzintervalle (**DIF-Plot**, siehe Abb. 62):

Im DIF-Plot werden Unterschiede der geschätzten Itemparameter durch horizontale Linien pro Item und Gruppe visualisiert, wobei die Gruppen durch die Liniengestaltung unterschieden werden. Die Längen der Linien entsprechen dem 95 %-Konfidenzintervall [...] des geschätzten Schwierigkeitsparameters in der jeweiligen Gruppe. Aus der Überlappung der Linien pro Item kann daher auch bei diesem Diagramm ein Rückschluss auf allfällige [sic!] signifikante Unterschiede der jeweiligen Itemschwierigkeit zwischen den Gruppen gezogen werden.“ (Koller et al., 2012, S. 86)

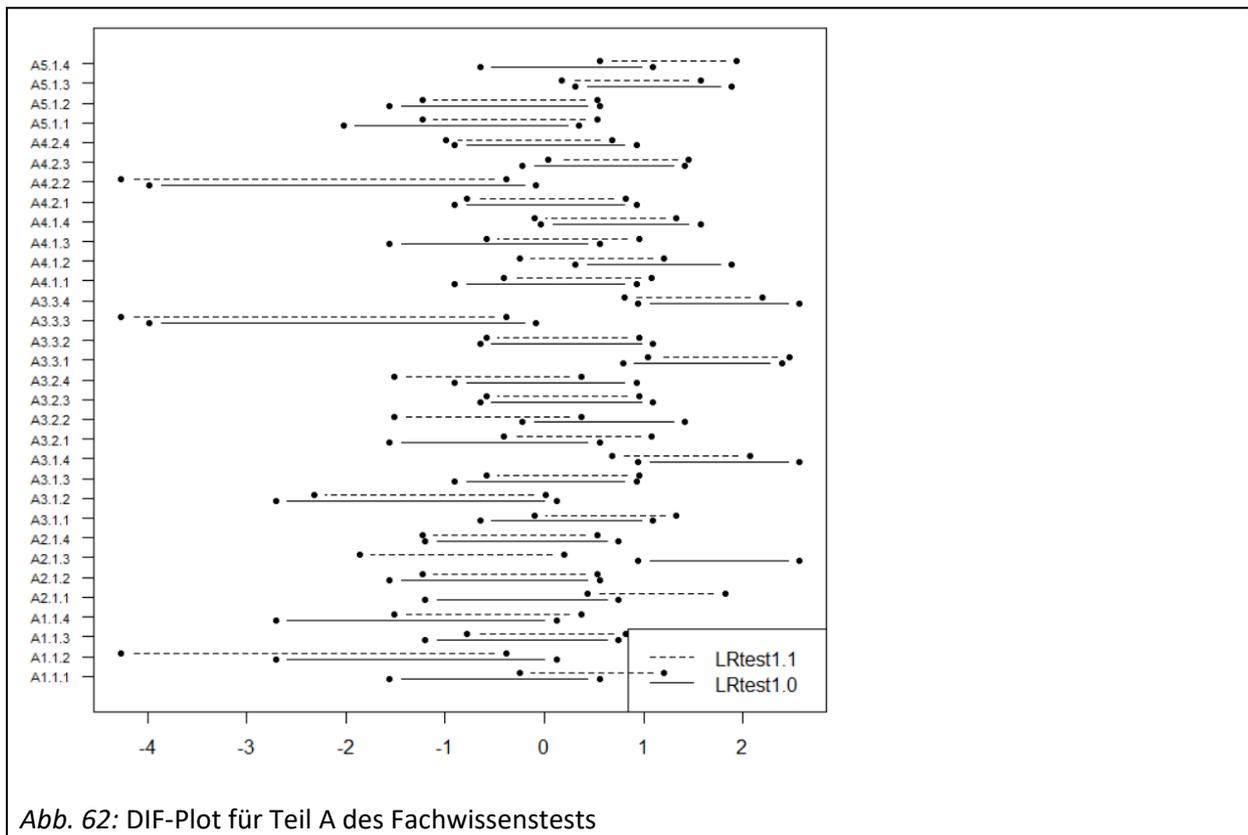


Abb. 62: DIF-Plot für Teil A des Fachwissenstests

Rechnerisch kann die graphische Feststellung auffälliger Items auch mit Hilfe des itemspezifischen **Waldtests** bestätigt werden, bei dem die Hypothese überprüft wird, ob sich die Itemparameter zwischen Subpopulationen unterscheiden und ob dichotome Items in einem Modell signifikant sind, d. h. ob sie zur Differenziertheit eines Modells beitragen oder eliminiert werden können, weil sie allenfalls vernachlässigbare Auswirkungen auf das Messergebnis haben (Engle, 2010). Die Teststatistiken unterscheiden sich vom bedingten Likelihood-Quotienten-Test. Anhand des errechneten p-Werts (als Überschreitungswahrscheinlichkeit) kleiner 0,05 kann ein Item als nicht-modellkonform identifiziert werden. Dabei gilt aber zu berücksichtigen, dass vor allem bei der Prüfung vieler Items das (nominale) Signifikanzniveau nicht eingehalten wird, sondern die Fehlerwahrscheinlichkeit viel höher sein kann als das 5 %-Niveau (Koller et al., 2012, S. 77). Koller et al. (2012, S. 176) empfiehlt daher zuerst die Bonferroni-Korrektur auf globalem Level, welche auf Itemebene noch durch die Anzahl der Items geteilt wird (siehe Kapitel 4.7.1.4.1). Eine solche Itemprüfung sollte nach Eid und Schmidt (2014, S. 188) jedoch erst durchgeführt werden, wenn der globale Waldtest, der auf der Ebene aller Items durchgeführt wird, signifikant ist. Um zu vermeiden, dass die Messung durch das Vorhandensein eines auffälligen Items verzerrt wird und andere Items fälschlicherweise einen signifikanten Wert erreichen, soll das am meisten auffallende Item zunächst allein ausgeschieden und dann das Raschmodell erneut berechnet werden. Mit allen weiterhin auffälligen Items soll in gleicher Weise verfahren werden, bis die Homogenität aller Items erreicht ist (Koller et al. (2012, S. 79)). Gemeinsam ist den Verfahren Andersen-Test, graphischer Modelltest und Waldtest, dass alle drei die Personenhomogenität als Prüfkriterium verwenden, d. h. es wird geprüft, ob der Test in zwei Stichproben dieselben Fähigkeiten oder Eigenschaften misst. Nachteilig daran ist, dass das Ergebnis davon abhängig ist, nach welchen Kriterien die Stichprobe in Teilstichproben aufgeteilt wird (Bühner, 2011, S. 539).

Eine weitere Voraussetzung liegt in dem oben schon genannten Kriterium der **Eindimensionalität**. Graphisch ist die Eindimensionalität im Rasch-Modell auch dadurch ausgedrückt, dass die X-Achse der ICC sowohl die Personen-Fähigkeit als auch die Aufgaben-Schwierigkeit (mit einer gemeinsamen, durch eine logarithmierte Transformation errechneten Einheit Logit) umfasst, was darauf hindeutet, dass Personen- und Aufgaben-Parameter auf einer gemeinsamen latenten Dimension liegen. Für den Fachwissenstest beispielweise bedeutet dies, dass auch nur das Fachwissen gemessen wird, ohne dass andere Konstrukte einen Einfluss auf das Messergebnis haben. Anhand von **Fit-Statistiken** für das Rasch-Modell lässt sich beurteilen, ob die Testaufgaben ein eindimensionales Konstrukt abbilden und wie gut die geschätzten Item- und Personenparameter in das Rasch-Modell passen („fitten“) bzw. welche Aufgaben des Testinstruments eine ausreichend hohe Raschkonformität aufweisen, um damit eine eindimensionale Rasch-Skala definieren zu können. Dabei wird die Abweichung zwischen der Lösungswahrscheinlichkeit, welche mit Hilfe der Rasch-geschätzten Parameter Aufgabenschwierigkeit und Personenfähigkeit berechnet wird, und der tatsächlichen Bearbeitung der Aufgabe untersucht. Zur Einschätzung des Item-Fits wird der Mean-Square (MNSQ) jeder Aufgabe in seinen beiden Ausprägungen als gewichteter Infit-Wert und als ungewichteter Outfit-Wert berechnet und mittels z-standardisierter t-Werte auf signifikante Abweichungen geprüft. „Mean square fit statistics indicate the size of (how much) misfit is revealed in the actual data. [...] The **standardized fit statistic** indicates ‘how likely’ is that amount of misfit. So, for (mis)fit: Mean squares indicate ‘how *much* misfit’; Z or t, however, indicate ‘how *likely* is the misfit“ (Bond & Fox, 2015, S. 67). Zur Interpretation der Zahlenwerte des MNSQ führt Neumann (2014, S. 363) an:

Hat eine Person nach dem Rasch-Modell eine 30-prozentige Wahrscheinlichkeit eine bestimmte Aufgabe zu lösen, dann ist der erwartete Wert 0,3. Gelingt dieser Person, die Aufgabe zu lösen, ist der beobachtete Wert 1. Die Abweichung beträgt 0,7. Die Berechnung der MNSQ-Werte erfolgt

nun für jede Aufgabe über alle Probanden. Der *Infit* berechnet sich als Quotient aus dem Mittelwert der quadrierten beobachteten Abweichungen und dem Mittelwert der erwarteten Abweichungen (d. h. der mittleren Varianz der Abweichungen).

Zusammenfassend sei dies nochmals an einem konkreten Beispiel erklärt. Wenn die Personenfähigkeit den gleichen logit-Wert erreicht wie die Aufgabenschwierigkeit (z. B. den Wert -2,5 logits), so bedeutet dies, dass diejenige Person die gestellte Aufgabe mit einer Lösungswahrscheinlichkeit von 50 Prozent lösen kann. Wenn die Personenfähigkeit größer ist als die Aufgabenschwierigkeit, so kann die Aufgabe von den betreffenden Personen mit einer über 50 Prozent liegenden Lösungswahrscheinlichkeit beantwortet werden. Wenn die Personenfähigkeit kleiner ist als die Aufgabenschwierigkeit, so wird die Lösungswahrscheinlichkeit kleiner als 50 Prozent.

Übereinstimmend mit Wang und Chen (2005, S. 377) soll folgendes Szenario vermieden werden: „If fit statistics are incorrect, a missfitting item (person) may not be located correctly or a good item (person) may be identified incorrectly as a missfitting item (person).“ Der Infit-Mean-Square-Wert (Infit-MNSQ) gibt dabei Auskunft über modellkonforme Antwortmuster im personenbezogenen mittleren Schwierigkeitsbereich (nahe des Null-Logits), der Outfit-Mean-Square-Wert (Outfit-MNSQ) berechnet sich mit Items, die personenspezifisch sehr schwierig oder sehr leicht (fern des Null-Logits) sind. Da der Infit-Mean-Square-Wert (Infit-MNSQ) sensibel auf Abweichungen zu modellkonformen Antwortmustern reagiert, dient er meist als Hauptindikator zur Interpretation der Itemqualität (Bond & Fox, 2015, S. 67). Der Outfit-Mean-Square-Wert (Outfit-MNSQ) hingegen lässt Items erkennen, die personenbezogen entweder extrem gering oder stark vom mittleren Schwierigkeitsgrad abweichen. Generell beruhen beide Itemfitparameter auf Chi-Quadrat-Statistiken, die durch ihre Freiheitsgrade dividiert wurden. Ihr Erwartungswert beträgt 1 und ihre Spannweite von 0 bis $+\infty$. Werte größer 1 kennzeichnen Items mit großer Variationsbreite (Underfit), es ist demnach mehr Varianz in den Daten vorhanden, als das Rasch-Modell auf Grundlage der bestimmten Personenfähigkeiten für diese Aufgabe erwarten würde. Dieser Underfit resultiert daraus, dass eine zu geringe Abhängigkeit des Antwortverhaltens von der Personenfähigkeit besteht, indem entweder sehr fähige Personen entgegen der Erwartung eine leichte Aufgabe nicht lösen können oder weniger fähige Personen eine sehr schwere Aufgabe lösen. Werte kleiner 1 kennzeichnen Items mit allzu strenger (deterministischer) Antwortstruktur (Overfit), bei denen es zu wenigen unerwarteten Antworten kommt, die Daten demnach überraschend gut ins Modell passen.³⁹ Der Overfit kann auf Verletzung der lokalen stochastischen Unabhängigkeit hindeuten, weil das Verhalten unerwartet präzise durch das Antwortverhalten im Rest der Aufgaben beschrieben werden kann (Bond & Fox, 2015, S. 239–241). Wilson (2005, S. 128f.) bemerkt hierzu: „items with [...] mean square greater than 1 [...] are those that contribute less toward the overall estimation of the latent variable, and hence are the ones that are most problematic in general.“ Linacre (2002, S. 878) stellt fest, dass Werte nahe 1 auf eine sehr hohe Messgenauigkeit hindeuten, und hält einen Wertebereich zwischen 0,5 und 1,5 als „productive for measurement“. Jedoch findet man in der Literatur auch strengere Cutoff-Werte für die Eliminierung von Items, was darauf hindeutet, dass noch kein allgemeiner Konsens über Ober- und Untergrenzen für die Abweichungen vom erwarteten Mean-Square besteht. Wang und Chen (2005) fanden heraus, dass die Stichprobengröße erheblichen Einfluss auf die Fitwerte hat, weshalb sie vorschlugen, die Grenze anhand der Anzahl der Testpersonen nach

³⁹ Ein Fitwert von 1,3 indiziert, dass es zu 30 Prozent mehr Variation zwischen den beobachteten und den erwarteten Antwortmustern gegenüber einer perfekten Passung zum Modell kommen kann. Ein Fitwert von 0,78 ($1 - 0,22 = 0,78$) bedeutet, dass es zu einer um 22 Prozent geringeren Variation gegenüber der perfekten Modellpassung kommt (Bond & Fox, 2015).

den Formeln $[1 \pm \frac{2}{\sqrt{N}}]$ für den Infit und $[1 \pm \frac{6}{\sqrt{N}}]$ für den Outfit zu berechnen. Für diese Studie würde dies bedeuten, einen Grenzbereich für den Infit von $1 \pm 0,26$ verwenden zu dürfen. Es wird aber trotzdem das strengere Intervall von 0,8–1,2 gewählt. Während Einigkeit darüber herrscht, Fit-Indizes deutlich über 1 in jedem Fall zu verwerfen, weil sie zur Messung nur wenig beitragen, wird die Eliminierung zu rigider Items durchaus kritisch betrachtet (Matschinger, o. A. d. J.).

Der t-Wert gibt auf einem Signifikanzniveau von 95 Prozent die Aussagekraft des Mean-Squares bezüglich der einzelnen Items an. Liegt der Infit- oder Outfit-t-Wert in einem Intervallbereich von -2,0 bis +2,0 bei Stichprobengrößen zwischen 30 und 300, so darf angenommen werden, dass die zugehörigen MNSQ-Werte nicht signifikant vom Erwartungswert 1 abweichen. t-Werte über 2,0 weisen einen Underfit auf und deuten auf eine zu geringe Messgenauigkeit der Items hin, während t-Werte unter -2,0 eine zu hohe Messgenauigkeit (Overfit) vermuten lassen, so dass Fehleinschätzungen nur mit einer sehr geringen Wahrscheinlichkeit auftreten (Bond und Fox (2015, S. 53) bezeichnen solche Ergebnisse als „too good to be true“)⁴⁰. Wilson (2005, S. 129) empfiehlt, die Eliminierung eines Items (als misfit) nur vorzunehmen, wenn sowohl die MNSQ-Werte als auch die t-Werte aus ihrem jeweiligen Intervall fallen. Rost (2004) sieht vor allem t-Werte über -1,96 wegen ihrer Messgenauigkeit als problematisch an und empfiehlt, entsprechende Items zu eliminieren, während t-Werte über +1,96 im Itempool belassen werden können. Mit Bond und Fox (2015, S. 67) bleibt somit abschließend festzuhalten: „making ‘fit’ decisions based on considering these four indices (Infit- und Outfit-MNSQ und Infit- und Outfit-t-Werte, Anm. d. Autors), is quite a balancing act; it might be easy to accept one of these indicators as more important for a particular purpose, but none is unimportant“.

4.7.1.5 Differential Item Functioning (DIF)

Die Überprüfung des **Differential Item Functioning (DIF)**⁴¹ stellt eine weitere Methode zur Validierung der Testitems und des gesamten Rasch-Modells dar und legt offen, ob diese Skala für alle Personen einer Stichprobe in gleicher Weise funktioniert. Dabei wird untersucht, ob sich die Aufgabenschwierigkeit einzelner Items zwischen zwei Gruppen mit unterschiedlicher Merkmalsausprägung (z. B. bezüglich des Geschlechts) unterscheidet. Geht man von der Grundannahme des Rasch-Modells aus, dass bei zunehmender Personenfähigkeit auch die Itemfunktion monoton ansteigt, so sind bei vorhandenem DIF die Itemfunktionen für unterschiedliche Subgruppen mit angenommener gleicher Personenfähigkeit verschieden. Statistisch wird bei den Subgruppen die Anordnung der Aufgaben auf der Logit-skala entsprechend ihrer Schwierigkeit unter Kontrolle der Personenparameter verglichen und der DIF-Wert als Differenz der entsprechenden Logit-Werte der Aufgabenparameter bestimmt. Über das Verhältnis der Itemschwierigkeit zur Standardabweichung wird entschieden, ob und welche Items signifikant DIF aufweisen. Dies ist nach Bond und Fox (2015) der Fall, wenn die Abweichung größer ist als der einfache Wert der entsprechenden Standardabweichung: „When an item’s difficulty estimate location varies across subsamples by more than the modeled error, then [...] evidence of DIF exists“ (S. 105).⁴² Zugleich empfehlen aber die gleichen Autoren, dass das Rasch-Modell dennoch validiert werden kann, wenn nur wenige Items eines Tests DIF aufweisen, während Westers und Kelderman (1991, S. 107)

⁴⁰ Rost (2004) hält einen Toleranzbereich zwischen -1,96 und +1,96 für die t-Werte als vertretbar.

⁴¹ Bond und Fox (2015, S. 109) bezeichnet DIF als „failures of invariance“, die auf potentielle Probleme mit dem Messinstrument aufmerksam machen sollen.

⁴² Scheuneman und Subhiyah (1998) sprechen für ihren 250-Items-Medizinertest mit über 400 Testpersonen schon bei einer Differenzabweichung größer 0,5 Logits von DIF.

ausführen: „A given [...] test consisting of many items with significant DIF may be unfair for certain subgroups, and it is important to identify these items, so that they can be improved or deleted from the test.“

Failures of invariance können auch mittels der theoretischen ICCs erklärt werden, indem die erwartete (logistische) Leistungskurve mit den tatsächlichen Ergebnissen verglichen wird. In Abb. 63 folgen in der oberen Graphik die tatsächlich erreichten Werte der Jungen (durchgezogene Linie) und der Mädchen (gestrichelte Linie) relativ zuverlässig der erwarteten ICC-Kurve, während in der unteren Graphik die beiden Leistungskurven einen recht unterschiedlichen Verlauf nehmen. Es ist eindeutig zu sehen, dass vor allem im Bereich geringer Personenfähigkeit die Mädchen deutlich schlechtere Werte erzielen als die Jungen, deren Kurve permanent oberhalb der ICC-Kurve verläuft. Aufgrund dieser graphischen Analyse ist anzunehmen, dass in Fall a) kein DIF vorliegt, während Fall b) auf gender-DIF hindeutet.

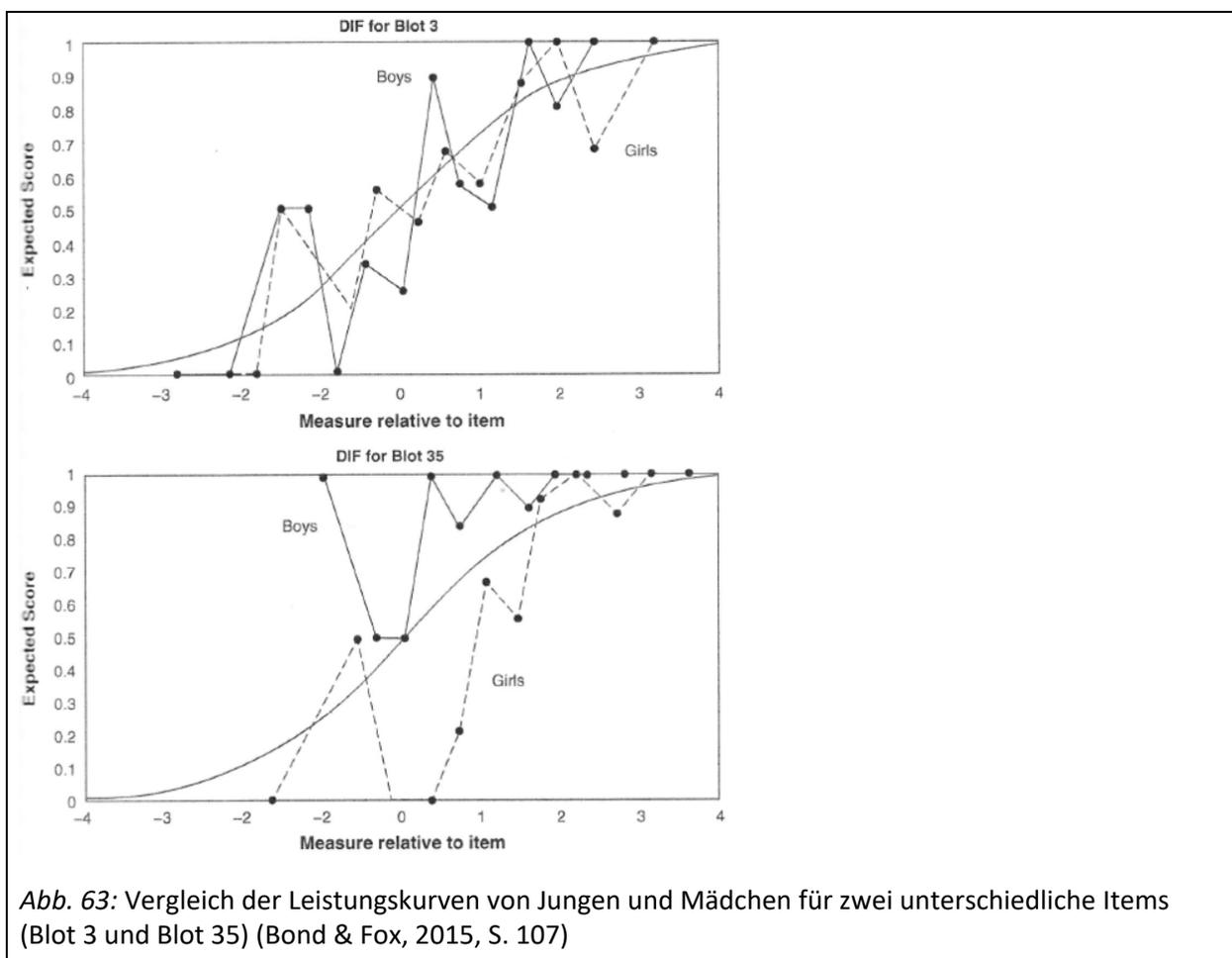


Abb. 63: Vergleich der Leistungskurven von Jungen und Mädchen für zwei unterschiedliche Items (Blot 3 und Blot 35) (Bond & Fox, 2015, S. 107)

Zur Feststellung von DIF können im Grunde die gleichen Verfahren angewendet werden, die auch schon für die Prüfung der Modellgültigkeit herangezogen wurden und nach Subgruppen unterscheiden. Der itemspezifische Waldtest ist in der Lage, DIF bei einzelnen Items aufzudecken, während andere Verfahren die globale Goodness-of-fit für das gesamte Rasch-Modell überprüfen, wie dies beim Likelihood-Ratio-Test nach Andersen der Fall ist. Der Vorteil dieser Testverfahren liegt darin, dass die Ergebnisse direkt dahingehend interpretiert werden können, welche Items von welcher Gruppe leichter oder schwieriger zu lösen sind. Dies kann ein Hinweis sein auf die Ursachen für das Auftreten von DIF und wie es eliminiert oder in zukünftigen Tests vermieden werden kann. Der Nachteil besteht allerdings darin, dass nur bestimmte, vom Forscher festgelegte Teilgruppen (z. B. nach den Splitkriterien

Geschlecht oder Alter) untersucht werden und DIF aufgrund anderer Gruppenunterscheidungsmerkmale unentdeckt bleibt (Strobl, Kopf & Zeileis, 2015, S. 290). Will man diese Einschränkung vermeiden, so ist das Mixed-Rasch-Modell (oder Latent Class (Mixture) Approach) von Rost (1990) anzuwenden, bei dem die Unterschiede der Itemparameter für alle möglichen Gruppen der Subjekte und unabhängig von den Personenmerkmalen untersucht werden. Nachteilig an diesem Modell ist die Tatsache, dass die Ergebnisse nicht leicht interpretierbar sind, sondern in einem zweiten Schritt erst mit den Personenmerkmalen in Verbindung gebracht werden müssen. Strobl et al. (2015) schlagen daher ein statistisches Verfahren (mithilfe des R-package *psychotree*) für die Identifizierung von DIF vor, das als Kompromiss aus den beiden bereits genannten Methoden konzipiert wurde, indem zwar alle vordefinierte Gruppen auf ihre Unterschiede, aber zugleich auch weitere potentielle Ursachen von DIF untersucht werden. In dieser Studie werden die Subgruppen mit dem Splitkriterium Geschlecht zur Identifizierung von DIF hauptsächlich mit dem graphischen Testverfahren und dem itemspezifischen Waldtest untersucht.

4.7.1.6 Trennschärfe

Die Trennschärfe eines Items gilt als Maß dafür, wie gut ein Item zwischen Personen mit niedriger und hoher Merkmalsausprägung trennen kann. Bortz und Döring (2009, S. 219) konkretisieren: „Der Trennschärfe eines Items ist zu entnehmen, wie gut das gesamte Testergebnis aufgrund der Beantwortung eines Items vorhersagbar ist.“ Die Trennschärfe muss daher für jedes Item einzeln berechnet werden und wird durch den Trennschärfekoeffizienten r_{it} dargestellt. Dieser gibt an, „wie groß der korrelative Zusammenhang der Itemwerte x_{vi} der Probanden mit den Testwerten x_v der Probanden ist“ (Kelava & Moosbrugger, 2012, S. 84). Die Trennschärfe entspricht daher einer Produkt-Moment-Korrelation, die auch als Pearson-Korrelation oder als bivariate Korrelation bezeichnet wird (Bühner, 2011, S. 172). Der Testwert x_v kann dabei entweder über die Zeilensumme sämtlicher Itemwerte x_{vi} eines Probanden oder als Summenwert ohne das betreffende Item bestimmt werden. Letztere Methode empfiehlt sich vor allem für die Trennschärfebestimmung bei wenigen Items (sog. *part-whole-correction* $r_{it(i)}$), um die Trennschärfe nicht zu überschätzen (Kelava & Moosbrugger, 2012, S. 84f.).

Das berechnete Korrelationsmaß r_{it} kann Werte zwischen 1 und -1 annehmen. Werte nahe bei 1 lassen die Schlussfolgerung zu, dass das Item von Probanden mit hoher Personeneigenschaft zuverlässiger gelöst wird als von Probanden mit niedriger Personeneigenschaft. Zugleich kann man davon ausgehen, dass die einzelnen Items sehr ähnlich differenzieren wie der Gesamtttest. Dabei gelten Trennschärfen im Bereich von 0,4 bis 0,7 als gute Trennschärfen. Bühner (2011, S. 81) beurteilt Trennschärfewerte über 0,5 als hoch, zwischen 0,3 und 0,5 als mittel und unter 0,3 als niedrig. Liegt r_{it} nahe 0, so lässt sich kein Zusammenhang zwischen dem Einzelitem und dem Gesamtttest herstellen, d. h. das Item ist ungeeignet, zwischen Probanden mit unterschiedlichen Personeneigenschaften zu differenzieren. Wenn die Trennschärfe nahe bei -1 liegt, dann wird das Item von Probanden mit niedriger Merkmalsausprägung zuverlässiger gelöst als von Probanden mit hoher Merkmalsausprägung. Die Gründe dafür können unterschiedlich sein (z. B. Mängel in der Instruktion oder bei der Itemformulierung, nicht-rückinvertierte Items). Ist das Item mit negativer Trennschärfe Teil eines Leistungstests, so deutet dies auf die Unbrauchbarkeit für die Messung der latenten Variable hin (Kelava & Moosbrugger, 2012, S. 86). Bühner (2011, S. 81) betont aber zugleich, dass „niedrige Trennschärfen nicht zwangsweise zur Entfernung der Items aus dem Test führen sollten“, weil diese Werte im Kontext der Gesamtuntersuchung und in ihrem Zusammenhang mit anderen Merkmalen wie der Schwierigkeit, der Reliabilität und der

Validität beurteilt werden müssen. Lienert und Raatz (1998, S. 122) stellen außerdem fest, dass bei der Itemselektion darauf geachtet werden muss, dass noch genügend leichte wie schwierige Aufgaben in der Testendform behalten werden sollen, weshalb manchmal auch Aufgaben mit geringer Trennschärfe in Kauf genommen werden.

Die Trennschärfe hängt demnach eng mit der Itemschwierigkeit zusammen. Je schwerer die Aufgabe ist, desto geringer fällt r_{it} aus (Bodeneffekt). Aber auch wenn die Aufgabe so leicht ist, dass annähernd jeder Proband dieses Item richtig löst, resultiert daraus eine geringe Trennschärfe für dieses Item (Deckeneffekt). Daraus folgt, dass Items mit mittlerer Schwierigkeit zwischen Probanden, die ein Item lösen („Löser“), und Probanden, die ein Item nicht lösen („Nicht-Löser“), am besten differenzieren oder trennen (Bühner, 2011, S. 527). Grundsätzlich sind Items mit hoher Trennschärfe wünschenswert, weil dadurch die Personeneigenschaften zwischen den Probanden differenziert erfasst werden.

Für intervallskalierte wie dichotome Daten gilt, dass eine hohe Trennschärfe im Allgemeinen durch eine hohe Itemvarianz begünstigt wird. Bei dichotomen Items kann die unkorrigierte Trennschärfe r_{it} einfacher als punktbiseriale Korrelation berechnet werden, welche in Formel 5 zur Berechnung der unkorrigierten Trennschärfe $r_{it(i)}$ eingeht:

$$r_{it} = \frac{\bar{x}_{v1} - \bar{x}_{v0}}{SD(x)} \sqrt{\frac{n_0 n_1}{n(n-1)}} \gg r_{it(i)} = \frac{r_{it} SD(x) - \sqrt{p_i(1-p_i)}}{\sqrt{SD(x)^2 + p_i(1-p_i) - 2r_{it} SD(x) \sqrt{p_i(1-p_i)}}} \quad (\text{Formel 5})$$

x_{v0} und x_{v1} beschreiben dabei die Mittelwerte der Probanden, die in Item i entweder eine 0 oder eine 1 als Antwort hatten. n_0 und n_1 ist die jeweilige Anzahl der Probanden, die in Item i eine 0 bzw. eine 1 als Antwort hatten. p_i stellt dabei die Lösungswahrscheinlichkeit für Item i dar, $SD(x)$ drückt die Standardabweichung der Testwerte aus. Die Standardabweichung des Items i wird in der Formel ausgedrückt durch seine Lösungswahrscheinlichkeit (p_i) und die entsprechende Gegenwahrscheinlichkeit ($1-p_i$) (Kelava & Moosbrugger, 2012, S. 85). Bei der Ermittlung der Trennschärfe auf Basis des Rasch-Modells wird aber auch kritisch gesehen, dass das Rasch-Modell vereinfachend für alle Items die gleiche Trennschärfe annimmt (Haladyna, 2004).

4.7.1.7 Schätzung der Personenparameter

Die Personenparameter werden unter Anwendung der Maximum-Likelihood-Methode über die R-Funktion `person.parameter()` geschätzt. Der Output zeigt die vorliegenden Rohwerte sowie den dazugehörigen Personenfähigkeitsparameter (Estimate) (siehe Tab. 28). Über den Befehl `summary(pers)` lässt sich für jede Testperson der Fähigkeitsparameter einschließlich dem Konfidenzintervall abrufen, welcher später für die Korrelationsberechnungen benötigt wird.

Tab. 28: Personenparameter des Fachwissenstests

Raw Score	Estimate	Std.Error
3	-1.47233421	0.6913666
4	-1.03693396	0.6332246
5	-0.65893941	0.5992360
6	-0.31233889	0.5802609
7	0.01901255	0.5727667
8	0.34783977	0.5759065
9	0.68692011	0.5908805
10	1.05261311	0.6216327
11	1.47136277	0.6778133
12	1.99849715	0.7858962

Eine besonders aussagekräftige graphische Darstellung der Personenparameter und Itemschwierigkeiten erfolgt mittels der Person-Item-Map über den Befehl `plotPimap()`. Die Verteilung der Personenparameter θ wird über ein Histogramm im oberen Bereich ebenso ersichtlich wie die Stelle, an der sich ein Item auf der latenten Dimension befindet und sich somit die jeweilige Aufgabenschwierigkeit β befindet (siehe Abb. 64).

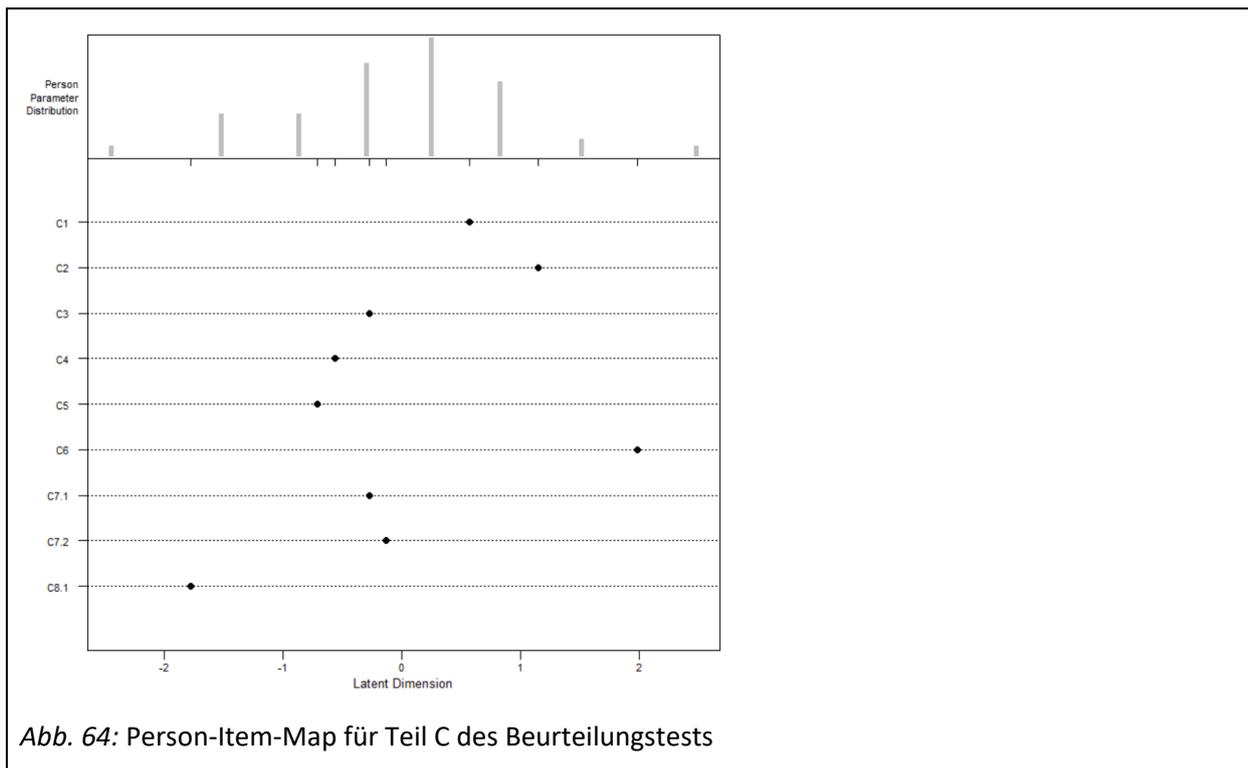


Abb. 64: Person-Item-Map für Teil C des Beurteilungstests

Wünschenswert wäre es, wenn die Items hinsichtlich ihrer Schwierigkeit über den ganzen Bereich der latenten Dimension verteilt wären. Bei einem sehr leichten Test sind die Personen eher auf der rechten Seite verteilt. Die Items hingegen werden für diese Personen eher zu leicht sein und liegen daher eher auf der linken Seite. Sie sind somit wenig geeignet, um die Fähigkeit dieser Personengruppe gut erfassen zu können. In Abb. 64 hingegen ist die Verteilung der Personeneigenschaften annähernd normalverteilt und die Items decken die ganze Breite der latenten Dimension ab.

Da der berechnete Personenwert nur eine geschätzte Größe ist, welche typischerweise nicht mit dem wahren Personenwert identisch ist, muss die Reliabilität für den Personenwert bestimmt werden. „Der Anteil der Varianz der wahren Personenwerte an der Varianz der geschätzten Personenwerte zeigt an, inwieweit die geschätzten Personenwerte wahre Personenunterschiede wiedergeben“ (Eid & Schmidt, 2014, S. 181). Dieses Verhältnis wird Personenseparierbarkeit (separation of person reliability) genannt und ist ein Maß dafür, wie wahrscheinlich es ist, dass die Rangfolge von Personen auf der latenten Personenfähigkeitsskala über ihre Messergebnisse beschrieben werden kann (Linacre, 2011, S. 618). Bei geringer Reliabilität macht es wenig Sinn, die Unterschiede in den geschätzten Personenwerten anhand anderer psychologischer Variablen erklären zu wollen. Insofern gibt der Wert für die Personenreliabilität an, mit welcher Wahrscheinlichkeit der Unterschied zwischen den geschätzten Personenwerten auf wahre Personenunterschiede zurückzuführen ist oder es sich nur um schätzfehlerbedingte Unterschiede handelt (Eid & Schmidt, 2014, S. 181f.). Die Rasch-Personenreliabilität ist in der Regel niedriger als der Cronbachs-Alpha-Wert, da die wahre Reliabilität im Rasch-Modell eher unterschätzt wird. Ihre Größe hängt außerdem von der Varianz der Personenfähigkeiten sowie von der Anzahl der Testaufgaben ab.

4.7.1.8 Das ordinale Rasch-Modell (Partial-Credit-Modell)

Der Nachteil des dichotomen Rasch-Modells für die Datenauswertung in dieser Studie besteht darin, dass einige Daten, die in kategorialer Form vorliegen, auf die Antwortmöglichkeiten richtig oder falsch reduziert werden müssen, um für die Modellrechnungen verwendet werden zu können. Dies bedeutet einen nicht immer vorteilhaften Verlust der Datenqualität, denn manchmal ist es wünschenswert auszudrücken, ob eine Antwort völlig, nur teilweise oder gar nicht mit der erwarteten Antwort eines oder mehrerer Experten übereinstimmt, wie dies beispielsweise beim Beurteilungstest dieser Studie (siehe Kapitel 4.5) der Fall ist.

Das Partial Credit Modell (PCM), welches von Masters (1982) entwickelt wurde und auch ordinales Rasch-Modell (ORM) genannt wird, kann diese kategorialen Variablen, deren Kategorien im Sinne der Zu- bzw. Abnahme der Merkmalsausprägung geordnet sind, verarbeiten. Die Bewertung von Leistungen mittels eines schulnotenähnlichen Systems, wie dies auch bei der Beurteilung der Eigenrealisation der Testpersonen angewendet wird (siehe Kapitel 5.2.3), kann ebenso auf diese Weise ausgewertet werden wie der Grad der Übereinstimmung mit einem Expertenurteil, welches ähnlich einer Likert-Skala operationalisiert werden kann. Dabei lässt sich die Verteilung einer kategorialen Variable wie bei dichotomen Variablen durch die Wahrscheinlichkeitsverteilung beschreiben, welche durch das Lagemaß (beispielsweise der Median oder bei Annahme gleicher Abstände zwischen den Kategorien der Mittelwert als Ausmaß der Schwierigkeit eines Items) sowie die Varianz (als Streuungsmaß der interindividuellen Unterschiede) gekennzeichnet ist. Das Streuungsmaß kann durch den relativen Informationsgehalt (entsprechend der relativen Häufigkeiten der Einzelkategorien) besser beschrieben werden als durch die Varianz, welche ihren Maximalwert erreicht, wenn die beiden Extremkategorien mit gleicher Wahrscheinlichkeit, die dazwischenliegenden Kategorien jedoch gar nicht besetzt wären. Dann wären die Daten jedoch wie dichotome Daten zu behandeln, da die mittleren Kategorien zu keiner Diskrimination der Personen beitragen würden (Eid & Schmidt, 2014, S. 224–228).

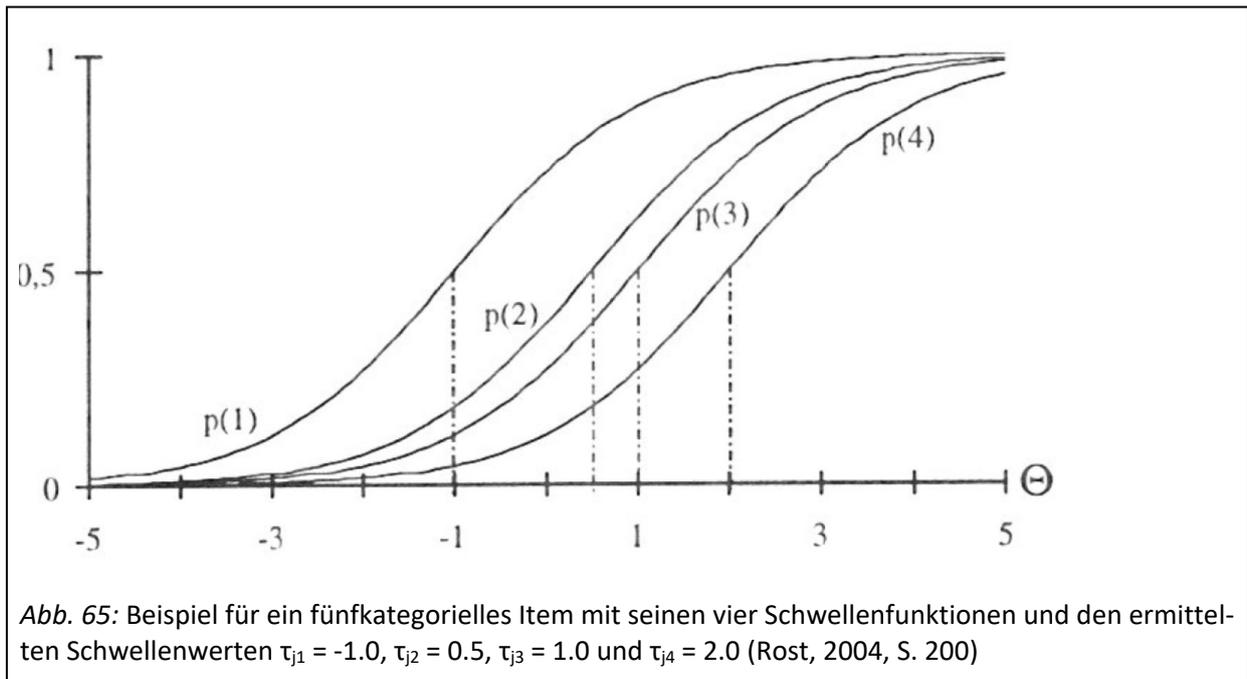
Da das Partial-Credit-Modell eine Erweiterung des dichotomen Rasch-Modells darstellt, sind viele Eigenschaften des Modells ähnlich zu bereits erwähnten Merkmalen (siehe Kapitel 4.7.1.2). Es bleibt beispielsweise die Annahme bestehen, dass der Summenwert einer Person eine suffiziente Statistik

für ihren latenten Personenwert darstellt, indem die Werte der Kategorien, welche die Person bei den verschiedenen Items ausgewählt hat, aufsummiert werden. Auch die Grundgleichung des Rasch-Modells (siehe Formel 6) kann auf Variablen mit geordneten Kategorien übertragen und in folgender Form angepasst werden:

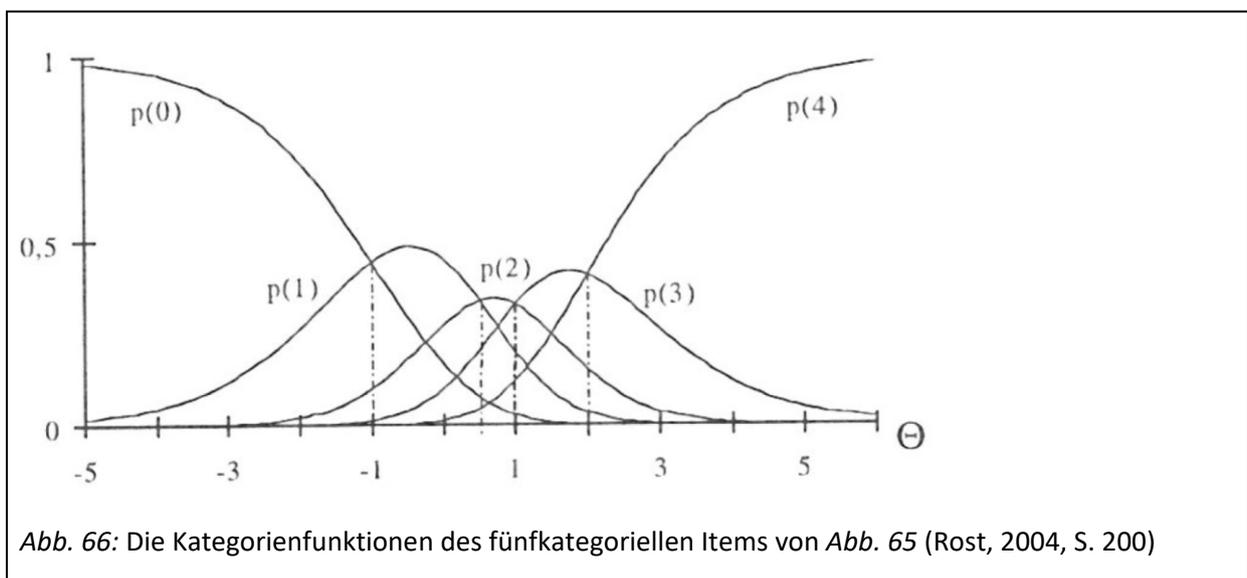
$$P(U_{ij} = 1/\theta_i) = \frac{e^{\theta_i - \tau_{jk}}}{1 + e^{\theta_i - \tau_{jk}}} \quad \text{für } x = 0, 1, \dots, k. \quad (\text{Formel 6})$$

Der Schwellenparameter τ_{jk} entspricht einer Konstante, die spezifisch für die beiden ersten Kategorien eines Items j ist, wobei der zweite Index x dem Umstand geschuldet ist, dass jedes Item mehrere Schwellen besitzt und jede Schwelle eine eigene Lokation (d. h. Lage auf dem latenten Kontinuum) hat. Der Begriff der Schwelle soll dabei suggerieren, „dass an diesem Punkt auf dem Kontinuum der Übergang von einer Kategorie zur anderen stattfindet, d. h. die Wahrscheinlichkeit der folgenden Kategorie zu antworten von diesem Punkt an größer wird als die Wahrscheinlichkeit, in der vorangegangenen Kategorie zu antworten“ (Rost, 2004, S. 198). Der ersten der k Kategorie wird der Wert $x = 0$ zugewiesen, der letzten Kategorie der Wert $x = k-1$, den Kategorien dazwischen aufsteigend natürliche Zahlen. Der Nenner erhält die bedingte Wahrscheinlichkeit, entweder die Kategorie x oder die darunterliegende Kategorie $x-1$ zu wählen. Formel 6 beinhaltet somit, dass die Schwellenwahrscheinlichkeit einer Person i bei Item j nicht nur von der Eigenschaftsausprägung dieser Person abhängt, sondern auch von der Schwierigkeit der Schwelle bei diesem Item (Rost, 2004, S. 200).

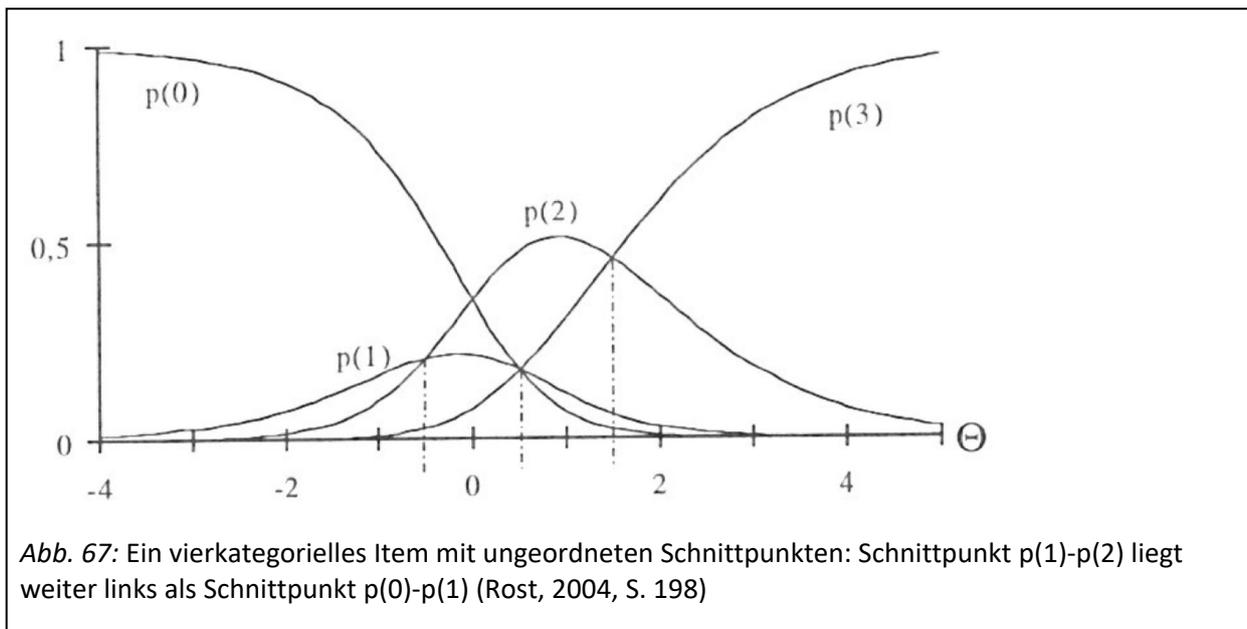
Die in Formel 6 errechnete Schwellenwahrscheinlichkeit beschreibt die Schwierigkeit (bzw. Wahrscheinlichkeit), die Schwelle von der unteren zur oberen Kategorie zu überschreiten. Ist der Personenwert θ_i gleich dem Schwellenparameter τ_{jk} , so ist die Wahrscheinlichkeit, eine der beiden Kategorien zu wählen, gleich. Ist $\theta_i > \tau_{jk}$, so präferiert die Person die obere Kategorie, im umgekehrten Fall die untere. Die Interpretation des Schwellenparameters entspricht dabei der Interpretation des Schwierigkeitsparameters β im Rasch-Modell und bezeichnet den Punkt auf der latenten Variable, ab dem die höhere Kategorie im Vergleich zur darunterliegenden Kategorie bevorzugt wird. Graphisch stellt sich diese Annahme so dar, dass für jede Schwelle eines Items eine logistische Funktion berechnet werden kann, in welcher der Schwellenparameter den Wert auf der latenten Variable beschreibt, bei dem die Schwellenwahrscheinlichkeit gleich 0,5 ist und die logistische Kurve ihren Wendepunkt (d. h. ihre maximale Steigung) hat (siehe Abb. 65). Je weiter links eine Kurve liegt, desto geringer ist der Personenwert, den man für die Überschreitung der entsprechenden Schwelle braucht. Im Idealfall sollte es so sein, dass die Kurve einer geringeren Kategorie weiter links liegt als die Kurve einer schwereren Kategorie, denn erst dann ist die Annahme erfüllt, dass es für das Überschreiten einer Schwelle zwischen zwei höheren Kategorien auch einer höheren Eigenschaftsausprägung bedarf (Eid & Schmidt, 2014, S. 232–234).



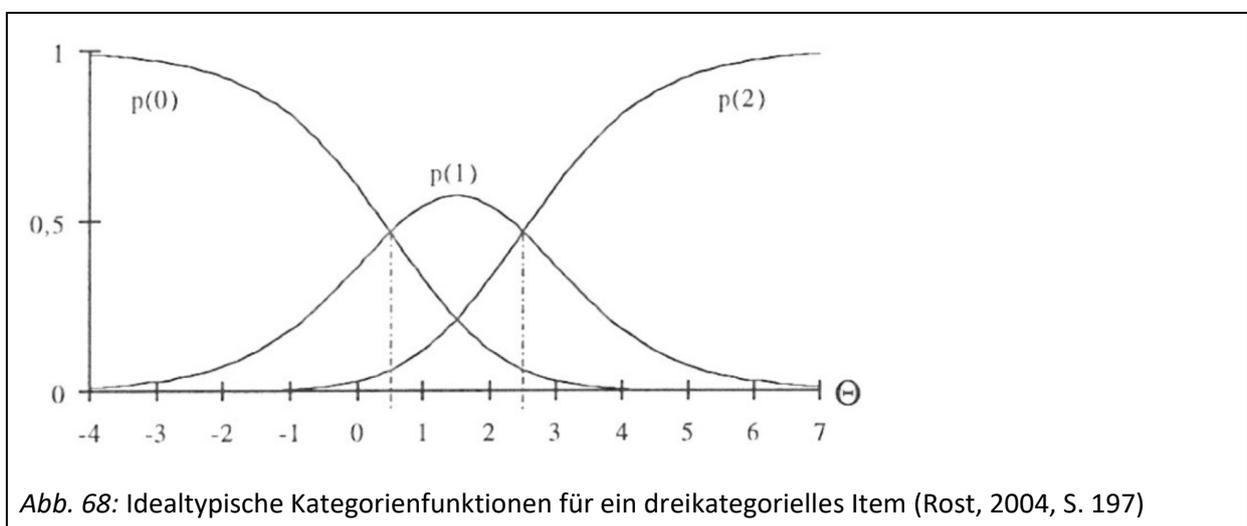
Rechnet man die pro Item vorliegenden Schwellenwahrscheinlichkeiten in Kategorienwahrscheinlichkeiten um, so ergibt sich eine Rekursionsformel, welche besagt, dass die Wahrscheinlichkeit der Kategorie x auf die Wahrscheinlichkeit der 0-ten Kategorie und auf eine Exponentialfunktion der entsprechenden Modellparameter zurückzuführen ist. Graphisch ergibt sich das Bild, dass die Lokationen der Schwellen auf dem Kontinuum (siehe Abb. 66) als Schnittpunkte der Kategorienfunktionen genau den Wendepunkten der Schwellenwahrscheinlichkeiten von Abb. 65 entsprechen. Die Extremkategorien haben eine monoton sinkende bzw. monoton steigende Kategorienfunktion, während die mittleren Antwortkategorien nicht-monoton eingipflig sind. Über die Abstände zwischen den Schnittpunkten der einzelnen Kurven kann die Größe bzw. Breite bestimmt werden, in welcher die jeweilige Antwortkategorie die relativ höchste Wahrscheinlichkeit besitzt. Somit wird jeder Antwortkategorie ein Abschnitt auf der zu messenden Personeneigenschaft (in Form der latenten Variable) zugewiesen, wodurch auch festgestellt werden kann, wenn Kategorien unterschiedlich groß sind (Rost, 2004, S. 200f.).



Ebenso kann die Ordnung der Kategorien überprüft werden, denn der Abschnitt für eine jeweils höhere Antwortkategorie muss auf der latenten Variable stets weiter rechts liegen, so dass eine höhere Personeneigenschaft für eine Antwort in dieser Kategorie erforderlich ist. Enthält eine Kategorie keinen eigenen Abschnitt, weil ihr Schnittpunkt mit der höheren Kategorie links vom Schnittpunkt mit der niedrigeren Kategorie liegt (siehe Abb. 67), so ist die präexperimentelle Hypothese der Ordnung der Antwortkategorien verletzt und die Ordinalskalenqualität der Itemantworten muss in Frage gestellt werden (Rost, 2004, S. 198).



In dieser Studie kommen entsprechend der drei Qualitätsstufen der Bewegungsbeurteilung dreikategorielle Items zur Anwendung, deren Kategorienfunktion idealerweise wie in Abb. 68 aussieht. Die Kurven verlaufen so, dass der Schnittpunkt zwischen der mittleren Kategorie ($p(2)$) und der Extremkategorie der Qualitätsstufe 1 ($p(0)$) weiter links liegt als der Schnittpunkt zwischen $p(1)$ und Qualitätsstufe 3 ($p(2)$). Zugleich sollte das Intervall für die mittlere Kategorie (zwischen den beiden Schnittpunkten) eine angemessene Größe haben, d. h. nicht zu weit oder zu eng zu sein, auch wenn sich diese Formulierung nicht in konkreten Zahlenwerten ausdrücken lässt.



Über die Kategorienfunktion lässt sich auch erkennen, wenn die Kategorien entgegen der präexperimentellen Hypothese vertauscht sind, d. h. der Personenwert für das Erreichen der Kategorie 2 weiter links liegt als für das Erreichen der Kategorie 1 (siehe Abb. 69, linker Graph für Item E1.2). Auf diese Weise können inhaltliche Widersprüche von Items aufgedeckt werden. Für das Item E1.2 würde das bedeuten, dass Testpersonen, welche das Hauptexpertenurteil gewählt haben, einen geringeren Personenwert aufweisen als Testpersonen, welche das Nahexpertenurteil gewählt haben. Dies kann beispielweise auf die Diskrepanz zwischen dem Expertenurteil und der Einschätzung der Testpersonen hinweisen. Item E1.6 (siehe Abb. 69, rechter Graph) hingegen zeigt, dass Personen mit einem Personenwert zwischen 0 und 2 mit einer hohen Wahrscheinlichkeit die mittlere Kategorie (Nahexpertenurteil) wählen, während Personen mit höheren Personenwerten das Hauptexpertenurteil wählen und Personen mit geringeren Personenwerten höchstwahrscheinlich ein Fehlurteil abgeben.

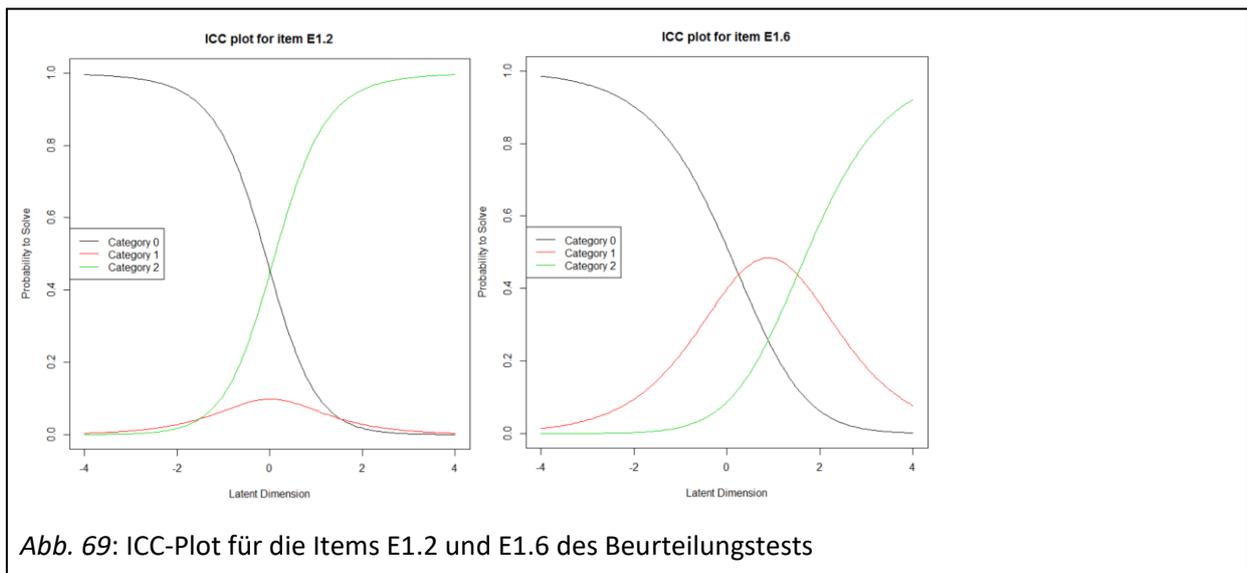


Abb. 69: ICC-Plot für die Items E1.2 und E1.6 des Beurteilungstests

Auch die Person-Item-Map des Partial-Credit-Modells lässt eine schnelle Interpretation der Datenanalyse zu. Es kann bestimmt werden, wo der Schwellenwert des Items im Bezug auf die Verteilung der Personenparameter (über das Histogramm) liegt, ob die Schwellenparameter geordnet oder ungeordnet sind und wie groß die entsprechenden Antwortkategorien sind (siehe Abb. 70).

Problematisch an der Datenauswertung nach dem Partial-Credit-Modell ist die Tatsache, dass jede Antwortkategorie im Datensatz vorkommen muss, damit die entsprechenden Schwellen berechnet werden können. Wenn eine Kategorie keine Fälle hat, so wird vor der R-Berechnung jede Kategorie dieses Items um eine Stufe nach unten bewertet, so dass die Rohpunktzahl dieses Items reduziert ist. Außerdem ist die Modellprüfung mit einer geringen Anzahl an Probanden nicht mittels quasi-exakter Tests möglich, sondern kann lediglich auf parametrische Weise erfolgen. Hierbei kann es durchaus zu einer falschen Annahme der Normalverteilung kommen, was wiederum die Modellprüfung verzerrt.

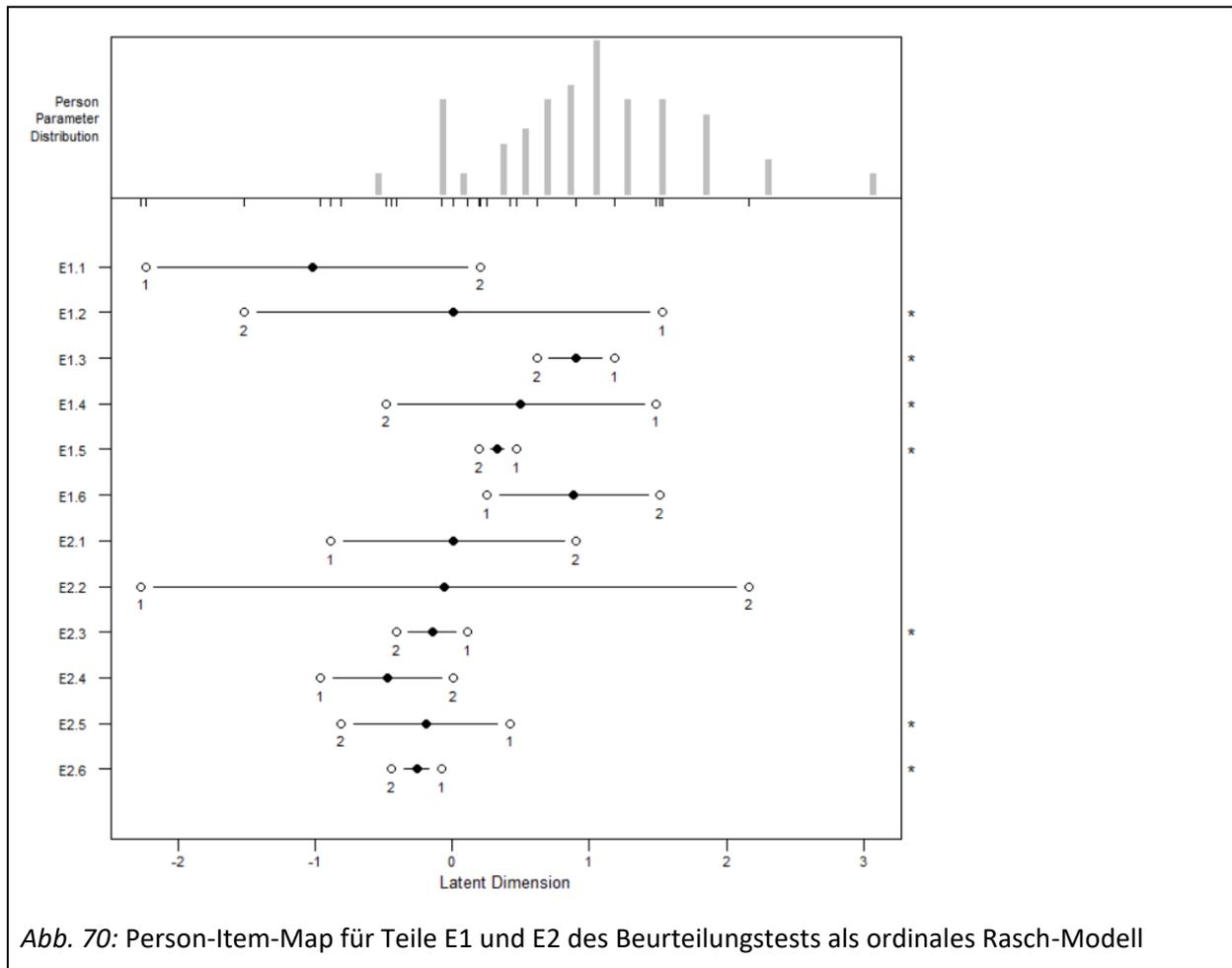


Abb. 70: Person-Item-Map für Teile E1 und E2 des Beurteilungstests als ordinales Rasch-Modell

Auch für das Partial-Credit-Modell müssen nämlich einige Annahmen erfüllt sein, um dieses zur Anwendung bringen zu können. Eindimensionalität und (lokale) stochastische Unabhängigkeit der Daten (Strobl, 2012) sind ebenso wie beim dichotomen auch für das ordinale Rasch-Modell eine Voraussetzung. Ebenso stellen die einzelnen Zeilenrandsummen eine suffiziente Statistik für den Fähigkeitsparameter θ und die entsprechenden Spaltenrandsummen eine suffiziente Statistik für den jeweiligen Schwellenparameter β dar. Insbesondere hat auch im PCM die Annahme der spezifischen Objektivität zu gelten und somit kein DIF vorzuliegen, indem sich die Bewältigung einer Schwelle bei Personen mit gleicher Personenfähigkeit nicht zwischen einzelnen Subgruppen unterscheidet. Zum konkreten Prüfen der Annahme der spezifischen Objektivität kann der Likelihood-Quotienten-Test nach Andersen oder auch der Waldtest (siehe Kapitel 4.7.1.4.2) verwendet werden.

Trotzdem eröffnet die Erweiterung des Rasch-Modells zum Partial-Credit-Modell mit der Möglichkeit, kategoriale Daten auf die gleiche Weise analysieren zu können und in Personenwerte umwandeln zu können, in dieser Studie interessante Möglichkeiten zum Datenvergleich. An einigen Stellen dieser Studie ist es sinnvoll, Ergebnisse sowohl mit dem dichotomen als auch mit dem ordinalen Rasch-Modell zu berechnen und Vergleiche anzustellen, ob die Ergebnisse so unterschiedlich sind, dass auch eine unterschiedliche Interpretation gerechtfertigt erscheint. Da in dieser Studie die Bewegungsqualität beispielsweise mittels dreier Qualitätsstufen eingeschätzt wird und eine Festlegung, welche Qualitätsstufe überhaupt die vermeintlich richtige ist, sogar bei erfahrenen Beurteilungsexperten uneinheitlich ausfällt, eröffnet das Partial-Credit-Modell die Möglichkeit, die Einschätzungen der Probanden nicht nur in richtig oder falsch zu operationalisieren, sondern entsprechend ihrer Übereinstimmung mit den

Expertenurteilen in drei Kategorien (0 = Fehlurteil, 1 = Nahexpertenurteil, 2 = Hauptexpertenurteil) zu erfassen. Somit kann eine Testperson, die weit vom Hauptexpertenurteil liegt, anders behandelt werden als eine Testperson, die nur knapp neben diesem liegt. Ob dieser Unterschied in der Datenanalyse möglicherweise auch die Ergebnisse der Beurteilungsleistung beeinflusst, kann dann näher untersucht werden.

4.7.2 Codierung

4.7.2.1 Auswertung nach dem dichotomen Rasch-Modell

Für die Datenanalyse des Beurteilungstests nach dem dichotomen Rasch-Modell (DRM) müssen die Antworten der Testpersonen in falsch (Wert 0) oder richtig (Wert 1) codiert werden. Dies bringt die Schwierigkeit mit sich, dass die Einschätzung der Bewegungsmerkmale in die drei Qualitätsstufen selbst unter den diese Studie unterstützenden Experten nicht immer einstimmig verläuft und somit eine Wertung in den Kategorien falsch oder richtig in manchen Fällen fragwürdig erscheint. Dennoch werden für die verschiedenen Teile des Beurteilungstests Kriterien entwickelt, nach denen eine dichotome Codierung in möglichst plausibler Weise vorgenommen wird, um die Daten nach dem DRM auswerten zu können.

Der einfachste Fall liegt vor, wenn sich die fünf Experten (observer a (obsa), observer b (obsb), obsc, obsd und obse) in ihrer Beurteilung einig sind und exakt die gleiche Einschätzung der Qualitätsstufen vornehmen, wie dies beispielsweise für fünf Items des Aufgabenteils A zutrifft (siehe

Tab. 29 für die Items A3 und A4). Wenn eine Testperson exakt die gleiche Einschätzung der Items wählt, wird das Urteil mit dem Wert 1 (Übereinstimmung mit dem Expertenurteil) codiert, in jedem anderen Fall mit dem Wert 0 (keine volle Übereinstimmung mit dem Expertenurteil).

Tab. 29: Experteneinschätzung der Items A3 und A4 mit Codierung der Antwortmöglichkeiten

		obsa	obsb	obsc	obsd	obse		QS1	QS2	QS3
A3	A3.1	2	2	2	2	2	>>	0	1	0
	A3.2	1	1	1	1	1	>>	1	0	0
	A3.3	3	3	3	3	3	>>	0	0	1
A4	A4.1	2	2	2	2	2	>>	0	1	0
	A4.2	3	3	3	3	3	>>	0	0	1
	A4.3	1	1	1	1	1	>>	1	0	0

Wenn die fünf Experten nicht zu einem einstimmigen Urteil kommen, muss nach Plausibilität entschieden werden. Als plausible Übereinstimmung wird gewertet, wenn sich mindestens drei Experten in ihrem Urteil einig sind und die übrigen Experten zugleich höchstens die nächstliegende Qualitätsstufe wählen, wie dies auf Item B4 in Tab. 30 zutrifft. In diesem Fall werden alle Testpersonen, welche ebenfalls die Qualitätsstufe 2 wählen, mit dem Wert 1 codiert, alle anderen mit dem Wert 0. Wenn sich jedoch beispielsweise drei Experten in ihrem Urteil der Qualitätsstufe 3 einig sind, ein weiterer Experte aber Qualitätsstufe 1 als Einschätzung wertet, wie dies bei Item B6 in Tab. 30 der Fall ist, so wird dies

als gravierende Uneinigkeit betrachtet und es muss darüber nachgedacht werden, das Item aus dem Test zu entfernen. Dieses Verfahren wird auf alle Items der Teile B, C, D und E angewendet, wenn eine dichotome Auswertung der Daten erfolgt.

Tab. 30: *Expertenurteile für Teil B des Beurteilungstests*

		obsa	obsb	obsc	obsd	obse
Bewegungstempo						
	B4	2	3	3	2	2
Bewegungsrhythmus						
	B6	3	2	1	3	3

4.7.2.2 Auswertung nach dem ordinalen Rasch-Modell

Für die Datenanalyse des Beurteilungstests nach dem ordinalen Rasch-Modell (ORM) erfolgt die Codierung der Antworten nach dem Grad der Übereinstimmung mit dem von den Experten getroffenen Bewegungsurteil. Als plausibelste Einschätzung wird die Qualitätsstufe bestimmt, welche von den meisten Experten gewählt wird (Hauptexpertenurteil, HEU), und mit dem Wert 2 codiert. Um auch auszudrücken, ob die Einschätzung der Testperson dem Urteil der Experten nah oder fern ist, werden die beiden anderen Qualitätsstufen mit dem Wert 1 (Nebenexpertenurteil, NEU) oder dem Wert 0 (Fehlurteil, FU) codiert. Dabei wird die Tatsache als ausreichend angesehen, dass mindestens ein Experte eine andere Einschätzung trifft als das Hauptexpertenurteil, um der Antwortmöglichkeit den Wert 1 zu verleihen. So kommt es bei den Items E1.1, E1.5, E2.2, E5.6 und E6.6 (siehe Tab. 31) vor, dass neben dem Hauptexpertenurteil die beiden anderen Qualitätsstufen als Nebenexpertenurteil zugelassen werden und es kein Fehlurteil gibt, weil jede der drei Qualitätsstufen von mindestens einem der Experten gewählt wurde. Wenn sich alle Experten im Hauptexpertenurteil einig sind, wird der nebenliegenden Antwort (falls HAU auf QS1 oder QS3 fällt) oder den beiden nebenliegenden Antworten, falls das Hauptexpertenurteil auf die Qualitätsstufen 2 fällt, der Wert 1 zugewiesen. Wenn die Experten neben dem Hauptexpertenurteil nur eine weitere Qualitätsstufe auswählen, so wird diese zum Nebenexpertenurteil, die andere Qualitätsstufe zum Fehlurteil deklariert, auch wenn das Hauptexpertenurteil der Qualitätsstufe 2 entspricht. Wenn eine Testperson das Antwortfeld für ‚keine Einschätzung möglich‘ gewählt hat, erhält sie immer den Wert 0. Auf diese Weise wird für die Beurteilungsleistung der Aufgaben von Teil E eine Matrix erstellt, welche aus den Werten 2 (Übereinstimmung mit dem Hauptexpertenurteil), 1 (Übereinstimmung mit dem Nebenexpertenurteil) und 0 (Fehlurteil oder keine Einschätzung oder fehlende Werte) besteht.

Dieses Codierverfahren wird auf alle Teile des Beurteilungstests bei der Datenauswertung nach dem ordinalen Rasch-Modell angewendet. Besonderheiten, die im Rahmen der Codierung wegen der verschiedenen Aufgabenstellungen nur bei einzelnen Teilen auftreten können, werden in den entsprechenden Kapiteln der Einzelteile genauer behandelt. Eine beispielhafte Auswertung der fünf Expertenurteile (observer a (obsa), oserver b (obsb), obsc, obsd und obse) der Einzelaufgaben der Beurteilungsaufgaben E1 und E2 mit Operationalisierung der Qualitätsstufen ist in Tab. 31 dargestellt.

Tab. 31: Codierung von Einzelitems entsprechend der Übereinstimmung mit den Einschätzungen der fünf Experten (observer a, b, c, d und e)

		obsa	obsb	obsc	obsd	obse		QS1	QS2	QS3
E1	E1.1	2	2	1	3	2	>>	1	2	1
	E1.2	2	2	2	3	2	>>	0	2	1
	E1.3	2	1	1	2	2	>>	1	2	0
	E1.4	2	2	1	2	2	>>	1	2	0
	E1.5	2	2	1	3	2	>>	1	2	1
	E1.6	1	2	2	2	1	>>	1	2	0
E2	E2.1	3	3	2	3	3	>>	0	1	2
	E2.2	3	2	1	3	3	>>	1	1	2
	E2.3	3	2	2	3	3	>>	0	1	2
	E2.4	3	3	2	3	3	>>	0	1	2
	E2.5	3	3	2	2	3	>>	0	1	2
	E2.6	3	3	3	3	3	>>	0	1	2

4.7.3 Normalverteilung

Eine Normalverteilung von Daten liegt vor, wenn sich die meisten Werte um den Mittelwert gruppieren, während die Häufigkeiten nach beiden Seiten hin gleichmäßig abfallen. Bei intervallskalierten Daten basieren viele statistische Berechnungsverfahren auf der Annahme, dass deren Werte eine solche Normalverteilung aufweisen. Um zu entscheiden, welche analytischen Tests angewendet werden dürfen, muss vorher festgestellt werden, ob die Werte hinreichend normalverteilt sind, also die gegebene Verteilung nicht signifikant von der Normalverteilung abweicht. Kann die Normalverteilung angenommen werden, so können parametrische Verfahren angewandt werden, im anderen Fall muss auf nicht-parametrische Berechnungen zurückgegriffen werden. Bei Unsicherheit bezüglich der Annahme der Normalverteilung werden in dieser Studie nicht-parametrische Berechnungsverfahren bevorzugt, weil sie auch auf normalverteilte Daten angewendet werden können, wenngleich in diesem Fall ein gewisser Verlust der Datenqualität sowie eine daraus folgende Berechnungsungenauigkeit akzeptiert werden müssen.

Um einen ersten Eindruck über die Werteverteilung einer Variable zu gewinnen, kann eine graphische Analyse des entsprechenden Datenhistogramms erfolgen. Der Vergleich der Balken (entsprechend der gewählten Balkenbreite) mit der idealtypisch eingezeichneten Normalverteilungskurve kann schon deutliche Hinweise auf das Vorliegen einer Normalverteilung liefern. Für die Variable BT_B_ORM (Beurteilungstest Teil B berechnet nach dem ordinalen Rach-Modell) kann eine solche Normalverteilung durchaus vermutet werden (Abb. 71). Eine strengere und in einem Zahlenwert ausgedrückte Prüfung nimmt der Kolmogorov-Smirnov-Test vor, der in SPSS innerhalb der explorativen Datenanalyse und in der Sammlung der nicht-parametrischen Tests zur Verfügung steht. Eine errechnete Irrtumswahrscheinlichkeit p kleiner als 0,05 zeigt an, dass die gegebene Verteilung signifikant von der Normalverteilung abweicht (Bühner & Ziegler, 2009, S. 174f.). Er ist besonders geeignet, wenn bei der zu testenden Variable eine begrenzte Anzahl von Kategorien vorliegt und durch die hohe Anzahl geteilter Rangplätze keine eindeutigen Rangfolgen auftreten (Bühner & Ziegler, 2009, S. 364f.).

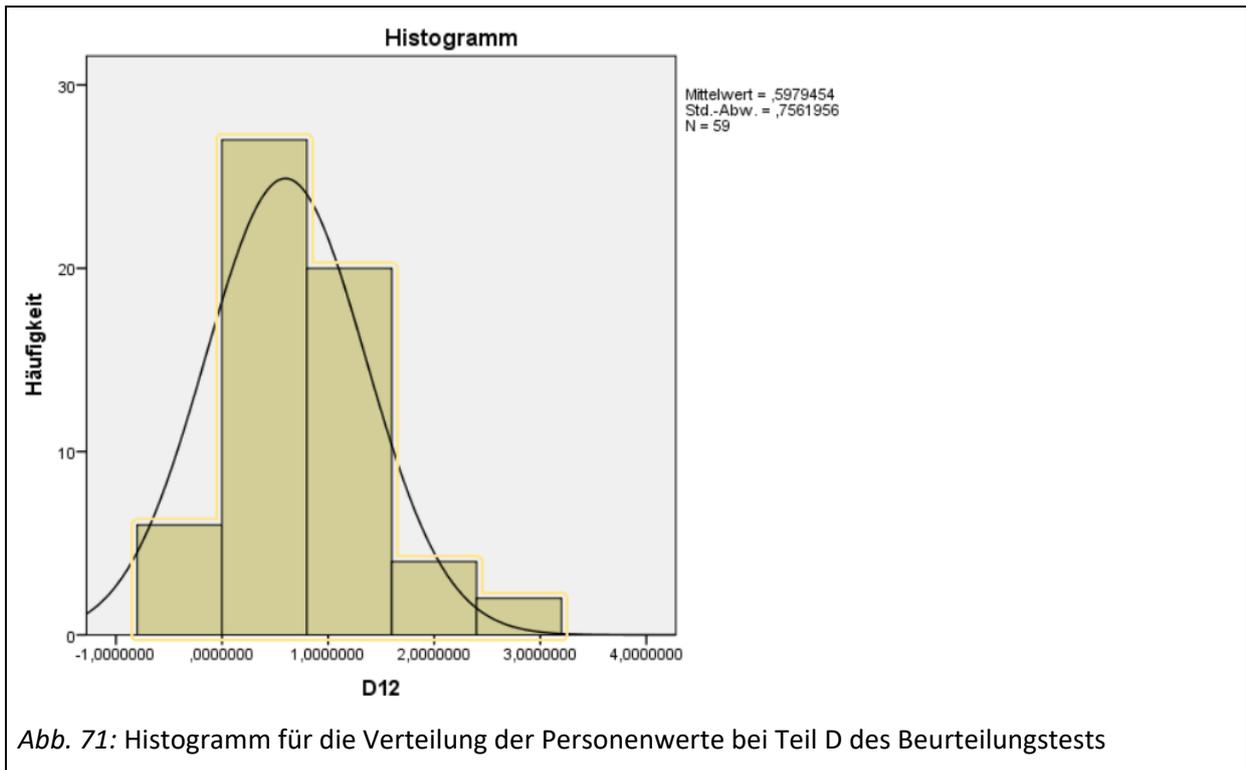


Abb. 71: Histogramm für die Verteilung der Personenwerte bei Teil D des Beurteilungstests

Eine weitere graphische Überprüfungsmöglichkeit bietet die Erstellung von Normalverteilungsplots (auch Q-Q-Diagramm genannt), in denen nach einer Transformation in z-Werte die beobachteten den unter Normalverteilung erwarteten Werten gegenübergestellt werden. Je weiter die beobachteten Werte von der Normalverteilungsgeraden abweichen, desto begründeter kann angenommen werden, dass keine Normalverteilung vorliegt (siehe Abb. 72, links). Im trendbereinigten Q-Q-Diagramm werden die Abweichungen zwischen beobachteten und erwarteten Werten in Abhängigkeit von den beobachteten Werten dargestellt (siehe Abb. 72, rechts), wobei im Normalverteilungsfall die Punkte auf einer horizontalen, durch den Nullpunkt verlaufenden Geraden liegen und ein deutliches Abweichungsmuster auf eine Abweichung von der Normalverteilung hindeutet (Bühner & Ziegler, 2009, S. 281f.).

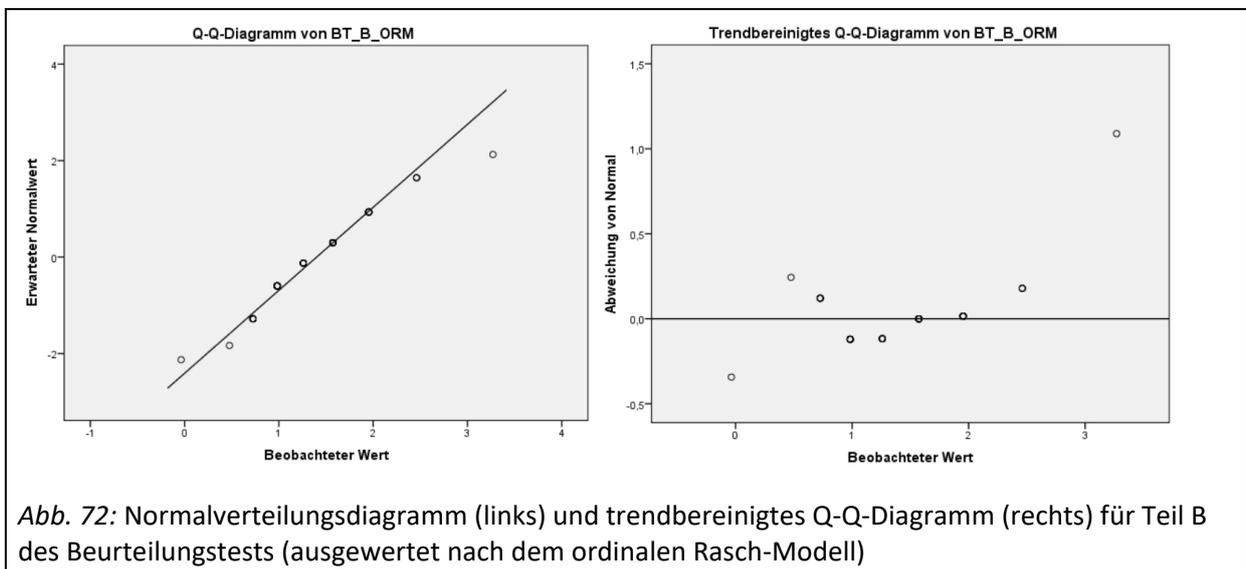
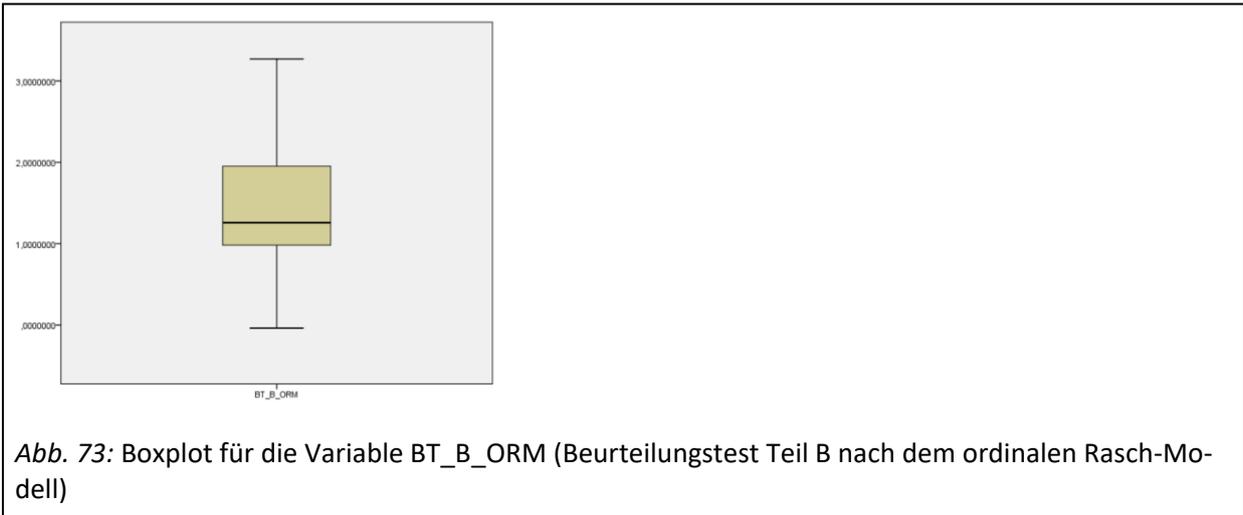
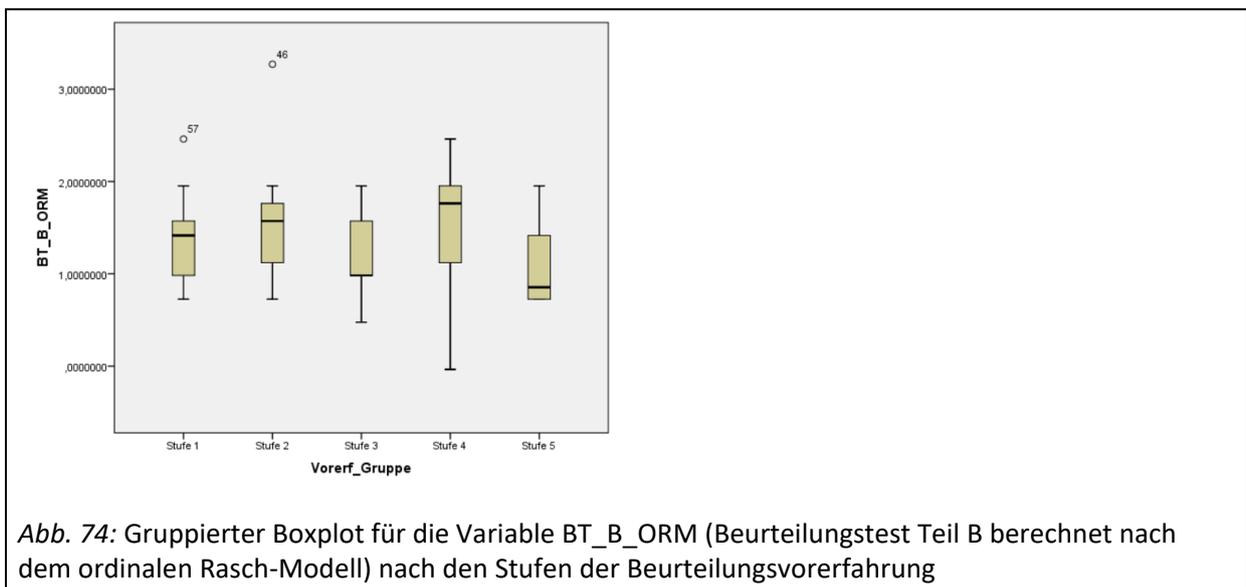


Abb. 72: Normalverteilungsdiagramm (links) und trendbereinigtes Q-Q-Diagramm (rechts) für Teil B des Beurteilungstests (ausgewertet nach dem ordinalen Rasch-Modell)

Zur Überprüfung von Extremwerten, welche die rechnerische Überprüfung der Normalverteilung verzerren können, kann die graphische Darstellung einer Variable mittels Boxplot hilfreich sein. Die innere, dickere Linie repräsentiert den Median, die beiden äußeren Linien begrenzen die Box vom ersten und dritten Quartil (25. bzw. 75. Perzentil). Auch der kleinste und der größte Wert werden über die außerhalb der Box liegenden Linien dargestellt. Extremwerte, die um mehr als drei Kastenlängen außerhalb der Box liegen, werden im Boxplot mit einem Stern markiert, solche um mehr als eineinhalb Kastenlängen außerhalb liegende mit einem Kreis (Bühner & Ziegler, 2009, S. 277). In Abb. 73 sieht man, dass die Variable BT_B_ORM (durchschnittlicher Personenfähigkeitswert für Teil B des Beurteilungstests, ausgewertet nach dem ordinalen Raschmodell) keine derartigen Ausreißer hat.



Der Boxplot kann für eine Variable auch entsprechend verschieden zu betrachtender Gruppen ausgegeben werden. In Abb. 74 ist wieder die Variable BT_B_ORM für die fünf verschiedenen Stufen der Beurteilungsvorerfahrung dargestellt. Es ist ersichtlich, dass Testperson 57 für Stufe 1 und Testperson 46 für Stufe 2 Extremwerte darstellen, die außerhalb des ersten bzw. dritten Quartils liegen und somit auch die Normalverteilungsannahme verzerren können. Bei Stufe 5 ist erkennbar, dass die Verteilung der Datenwerte unterhalb des Median deutlich geringer ist als oberhalb, was auch wieder Annahmen über die Verteilungsform zulässt.



4.7.4 Korrelations- und Regressionsanalyse

Die Überprüfung möglicher Zusammenhänge zwischen Variablen wird mit SPSS und R durchgeführt. Grundsätzlich muss beachtet werden, ob die Daten normalverteilt sind oder nicht. Bei normalverteilten Daten wird der Pearson's-Koeffizient verwendet, als nicht-parametrische Methode wird von Bühner und Ziegler (2009, S. 616) die Berechnung von Spearman's Rangkorrelation oder Kendall's Tau vorgeschlagen. Beide letzteren Verfahren beruhen auf den Rangplätzen der Testpersonen (d. h. der Proband mit dem niedrigsten Wert erhält Rangplatz 1, der mit dem zweitniedrigsten Rangplatz 2 usw.), wobei bei ordinalen Daten die erreichten Punkte in Rangplätze transformiert werden. Die beiden Berechnungsverfahren unterscheiden sich darin, dass für Spearman's Rangkorrelation die Rangplätze entweder gleichabständig sind oder intervallskalierte Rohwerte in Rangplätze überführt werden, während für Kendalls-tau diese Annahme nicht vorliegen muss. Bei gleichen Rängen werden nach Spearman die Rangplätze aufgeteilt (d. h. wenn zwei Personen gemeinsam die schlechteste Punktzahl erreicht haben, so belegen sie den Rang 1.5, während die nächsthöhere Person den Rang 3 belegt). Wenn die Ränge einer Person für zwei Variablen auf diese Weise vorliegen, fließt die quadrierte Differenz jeder Person zwischen den Rängen beider Variablen in die Berechnung der Rangkorrelation ein. Kendalls-tau hingegen ist weniger anfällig gegen Ausreißer und kann ordinale Daten genauer verarbeiten, indem nicht die erreichten Ränge selbst, sondern die Häufigkeiten der Rangplatzüberschreitungen und Rangplatzunterschreitungen im Vergleich zu den anderen Testpersonen ermittelt und als Rangsummendifferenz jeder Testperson in die Berechnung des Koeffizienten eingehen. Testpersonen mit gleichem Rangplatz werden bei Kendall's Tau nicht miteinander verglichen.

Die Datenauswertung über das Rasch-Modell entspricht sehr gut dem Prinzip der Rangkorrelation, weil die ermittelten Personenfähigkeitswerte nicht ordinal sind. Die Rasch-Berechnung ordnet jedem Rohpunktwert einen spezifischen Personenfähigkeitswert zu, wodurch letzterer nicht jeden beliebigen Zahlenwert annehmen kann. Somit entspricht der Personenfähigkeitswert einem bestimmten Rang, den die Testperson erreicht hat. Der Zahlenwert der Personenfähigkeit spielt zwar für die Interpretation der Personenfähigkeit selbst eine Rolle, nicht jedoch für den Vergleich der Personen zueinander. Es erscheint aufgrund der Datenbeschaffenheit gerechtfertigt, für Korrelationsberechnungen in dieser Studie Spearman's Rangkorrelation den Vorzug gegenüber Kendall's Tau zu geben.

Für die graphische Analyse müssen die Daten weiter aufbereitet werden, um Zusammenhänge zu erkennen. Dabei ist die Jitter-Funktion eine wichtige Möglichkeit, um der häufig vorkommenden Überlappung von Datenpunkten zu begegnen und die dadurch verborgene Datenstruktur zu erkennen (siehe Abb. 75). Auf diese Weise wird den Daten mithilfe von R ein natürliches Rauschen hinzugefügt (mit den Jitter-Werten 3 und 4). Dieses Problem tritt nicht nur bei der Verwendung einer Likert-Skala auf, sondern auch im Rahmen dieser Studie, weil die errechneten Personenwerte, die miteinander verglichen werden, je nach erreichter Rohpunktzahl ebenfalls feste Werte annehmen, die wie die Stufen einer Likert-Skala interpretiert werden können. Durch ein geringfügiges Jittern der Daten werden die Messwerte einer kontinuierlichen Skala angenähert und Tendenzen im Zusammenhang zwischen den untersuchten Variablen können leichter erkannt und interpretiert werden (Muenchen, 2011, S. 483).

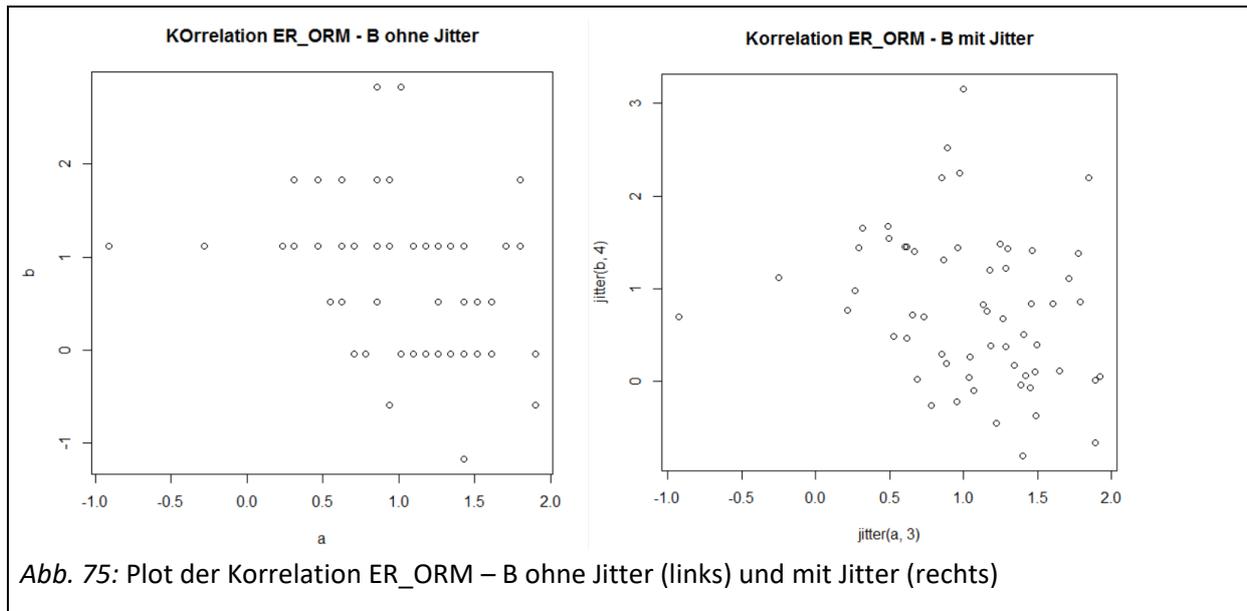


Abb. 75: Plot der Korrelation ER_ORM – B ohne Jitter (links) und mit Jitter (rechts)

Da eine Korrelationsberechnung noch keine Aussage über die Wirkungsrichtung der Variablen ermöglicht, soll bei auffallend hohen Korrelationskoeffizienten eine lineare Regressionsanalyse durchgeführt werden, um zu klären, ob die Unterschiede in einer abhängigen Variablen Y auf Unterschiede in einer unabhängigen Variablen X zurückgeführt werden können. Zur Interpretation, ob ein Korrelationskoeffizient als hoch zu bewerten ist, kann folgende Orientierungshilfe nach Cohen (1988) dienen: „Eine Korrelation zwischen $r = 0.10$ und $r = 0.30$ wird als gering bis moderat bezeichnet und eine Korrelation zwischen $r = 0.30$ und $r = 0.50$ als moderat bis groß“ (Bühner & Ziegler, 2009, S. 607).

Bei Vorliegen auffälliger Kriteriumskorrelationen (man spricht auch von Kriteriumsvalidität) soll in dieser Studie weiter untersucht werden, ob die Prädiktorvariablen Fachwissen, Eigenrealisation und Fähigkeit zum Ist-Soll-Wert-Vergleich oder zur Beurteilungsvorerfahrung Einfluss nehmen auf den erzielten Personenparameter einer Beurteilungsaufgabe (abhängige Variable). Zu diesem Zweck wird eine multiple lineare Regression berechnet, welche das Ziel hat, eine abhängige Variable durch mehrere unabhängige Variablen vorherzusagen. Um die Regressionsbeziehungen nachvollziehbar zu überprüfen, werden die Variablen Korrelationskoeffizienten größer 0,1 für das Regressionsmodell herangezogen. Dabei gilt es zu beachten, dass kein wichtiger Prädiktor in der Regressionsanalyse fehlt, weil dadurch die Schätzung der β -Gewichte und der Konstante verzerrt sein können und es zur Überschätzung der untersuchten Variablen kommt. Problematisch wird dies vor allem dann, wenn ein nicht berücksichtigter Prädiktor stark mit den anderen Prädiktoren korreliert, weil dieser die eigentlich gemeinsame Varianz mit dem Kriterium reduzieren würde und auch bei der Vorhersage des Kriteriums fehlt, wodurch die verbleibenden Gewichte einen Bias aufweisen. Deshalb muss vor Beginn der Regressionsanalyse überprüft werden, ob es auffällige Korrelationen zwischen den Prädiktoren gibt, um dann das Regressionsmodell entsprechend anzupassen. Zugleich muss der Tatsache Rechnung getragen werden, dass bei zu vielen miteinander korrelierende Prädiktoren die β -Gewichte sehr gering ausfallen und somit der Nachweis signifikanter Effekte schwierig wird. Daher erscheint es vor allem bei einer geringen Stichprobenzahl durchaus gerechtfertigt, eine begründete Auswahl der mit dem Kriterium korrelierenden Variablen zu treffen und die Interkorrelation der Prädiktoren nicht überzubewerten (Bühner & Ziegler, 2009, S. 679f.).

Dabei wird in SPSS zunächst die Standardmethode verwendet, bei der alle Prädiktoren gleichzeitig zur Vorhersage der abhängigen Variable betrachtet werden. Die Methoden der vorwärts- und rückwärtsgerichteten sowie schrittweisen Regression sind aufgrund zahlreicher Nachteile zwar laut Bühner und Ziegler (2009, S. 688) bedenklich, jedoch soll nach einem inhaltlich begründeten bzw. theoretisch fundierten Vorgehen mittels Rückwärtselimination überprüft werden, welche Prädiktoren letztendlich signifikant sind, um für diese eine entsprechende Regressionsanalyse durchzuführen.

4.7.5 Nicht-parametrischer Gruppenvergleich

Um den Zusammenhang zwischen den erreichten Personenwerten in den einzelnen Teilen des Beurteilungstests und der Vorerfahrung in der Bewegungsbeurteilung detaillierter zu überprüfen (siehe Kapitel 5.1.4), wird ein nicht-parametrischer, verteilungsfreier Gruppenvergleich durchgeführt, weil die Annahme der Normalverteilung für die vorliegenden Daten nicht aufrechterhalten werden kann (siehe z. B. Kapitel 5.1.1). Nicht-parametrische Berechnungsverfahren beruhen dabei nicht auf den Messwerten selbst, sondern auf einer gemeinsamen Rangreihe der Werte der zu untersuchenden Stichproben. Da es sich um mehr als zwei unabhängige Stichproben (nämlich die fünf Vorerfahrungsstufen) handelt, deren Unterschiede untersucht werden sollen, wird der **H-Test nach Kruskal und Wallis** durchgeführt, welcher eine Ausweitung des Mann-Whitney-U-Test zum Vergleich zweier unabhängiger Stichproben darstellt (Bühl, 2016, S. 373–375). Der H-Test prüft eine unterschiedliche Lage von Verteilungen (über die Medianwerte) und entspricht einer einfaktorischen Varianzanalyse für Rangziffern (Janssen & Laatz, 2017, S. 666). Dabei wird vorausgesetzt, dass das Merkmal (in diesem Fall die Beurteilungsleistung) in der Population stetig ist, die Stichproben voneinander unabhängig sind, die Messwerte innerhalb einer Gruppe voneinander unabhängig sind und die Verteilung des Merkmals sich zwischen den Subpopulationen nicht unterscheiden. Das Argument, dass die Varianzhomogenitätsannahme aufgrund von Ausreißern verletzt wird, reicht allein nicht aus, denn auch die Streuungen in den Subpopulationen müssen homogen sein. Wenn die Stichproben sehr klein sind (wie es in dieser Studie der Fall ist) und es in den Gruppen einzelne Ausreißer- oder Extremwerte gibt, so kann dies die Ursache für die Varianzheterogenität sein. „In diesem Fall ist die Anwendung des H-Tests eine Lösung, denn während einzelne Ausreißer und Extremwerte auf den Bedingungs Mittelwert (und die Bedingungsvarianz) einen großen Einfluss haben können, ist ihr Einfluss im Falle von Rangsummen wesentlich geringer“ (Eid, Gollwitzer & Schmitt, 2017, S. 440). Um den H-Test in dieser Studie durchführen zu können, werden die Teilnehmer entsprechend ihrer jeweiligen Zugehörigkeit zu den Stufen codiert und dann die Berechnung mit den entsprechenden Teil- und Gesamtergebnissen des Beurteilungstests als Variablen durchgeführt.

Das Ergebnis des Kruskal-Wallis-Tests kann durch den allgemeineren Mediantest bei unabhängigen Stichproben abgesichert werden, wenn die Messwerte der Testvariable unzuverlässig sein könnten. Auch der Jonckheere-Terpstra-Test, der bei Vorliegen eines vermuteten Trends angewendet werden soll (z. B. wenn mit der Zunahme der unabhängigen Variable (in unserem Fall der Vorerfahrungsstufe) auch eine Zunahme der abhängigen Variable (in unserem Fall der Beurteilungsleistung) vermutet wird), kommt in dieser Studie zum Einsatz, um die Ergebnisse des Kruskal-Wallis-Tests zu bestätigen oder zu widerlegen (Janssen & Laatz, 2017, S. 668–674).

Um weiterhin zu überprüfen, ob einzelne Gruppen sich paarweise voneinander unterscheiden, wird für jede Gruppenkombination ein entsprechender Mann-Whitney-U-Test mit SPSS (unter den nicht-parametrischen Tests über das alte Dialogfeld Tests bei zwei unabhängigen Stichproben) durchgeführt.

Um die Bedeutsamkeit signifikanter Ergebnisse zu beurteilen, wird die Effektstärke des Dunn-Bonferroni-Tests nach Formel 7 berechnet, welche der Effektstärke eines Rangsummentests entspricht:

$$r = \left| \frac{z}{\sqrt{N}} \right| \quad (\text{Formel 7}), \quad z = \text{z-Wert des Dunn-Bonferroni-Tests}, \quad N = \text{Stichprobengröße beider Gruppen}$$

Als Maß für die Einschätzung der Effektstärke erfolgt die Orientierung an Cohen (1992, S. 156), der bei $r = 0,1$ von einem schwachen, bei $r = 0,3$ von einem mittleren und bei $r = 0,5$ von einem starken Effekt spricht.

Auf das Problem des Mann-Whitney-U-Tests muss jedoch hingewiesen werden, welches in seiner Rangbindung bei Ties liegt, d. h. dass zwei Personen mit dem gleichen Personenwert auch den gleichen Rang erhalten und sich diesen teilen. Wenn also zwei nebeneinanderliegende Ränge (z. B. die Ränge 10 und 11) auf zwei Personen fallen, die den gleichen Personenwert haben, so erhalten beide als Rang den Mittelwert (hier: 10,5), wodurch eine Trennung der beiden Gruppen nicht mehr möglich ist und die Teststärke sinkt. Bei zu vielen Ties müsste ein Korrekturverfahren angewendet werden, welches aber für nicht-parametrische Testverfahren in SPSS nicht möglich ist.

5 Ergebnisse

Im folgenden Kapitel werden die Ergebnisse der Studie über die möglichen Einflussfaktoren auf die Beurteilungsleistung sportlicher Bewegungen und ihre gegenseitigen Zusammenhänge dargestellt. Dabei wird gleich in Kapitel 5.1 dargestellt, welche Korrelations- und Regressionsbeziehungen zwischen den Personenfähigkeitswerten der einzelnen Beurteilungsaufgaben und den gemessenen Variablen Fachwissen, Fähigkeit zum Ist-Soll-Wert-Vergleich, Eigenrealisation sowie Beurteilungsvorerfahrung festgestellt werden konnten. Die detaillierte Auswertung und Analyse der einzelnen Testverfahren, die in ihrem methodischen Aufbau in Kapitel 4 beschrieben wurden, erfolgt dann in Kapitel 5.2. Hier wird erklärt, wie die Personenparameter der einzelnen Variablen zustande kommen, wie sich die einzelnen Items bewährt haben und wie die Modellauswahl der statistisch angewendeten Verfahren geprüft wurde.

5.1 Zusammenhänge der möglichen Einflussfaktoren mit der Beurteilungsleistung

5.1.1 Messung der Beurteilungsleistung

Zunächst werden die allgemeinen Ergebnisse der Messung der Beurteilungsleistung behandelt, weil in den folgenden Kapiteln gleich auf die Zusammenhänge zwischen den Personenparametern der einzelnen Beurteilungsaufgaben und den Personenparametern der zu untersuchenden Einflussgrößen eingegangen wird. Um die Leistungsfähigkeit der einzelnen Testpersonen beim Beurteilen der Schlagwurf-bewegung zu messen, wurden die Personenfähigkeitswerte für die fünf Teile des Beurteilungstests nach dem Rasch-Modell ermittelt (siehe Kapitel 5.2.4). Betrachtet man den durchschnittlichen Personenparameter der Testgruppe für jeden Aufgabenteil, so nimmt dieser – abgesehen von Aufgabe C (Bewegungsvergleich) – kontinuierlich ab. Die Testpersonen erreichen mit zunehmender Komplexität der Beurteilungsaufgabe im Durchschnitt mit ihrer Personenfähigkeit eine geringere Lösungswahrscheinlichkeit, obwohl durch die Bearbeitung der vorherigen Testteile sicherlich ein nicht zu vernachlässigender Übungseffekt durch das Vertrautwerden mit den Bewegungsmerkmalen vorliegt. Testteil C (Bewegungsvergleich) ist dabei die am schwierigsten zu bewältigende Aufgabe. Diese Tendenzen treffen unabhängig von der Auswertungsmethode nach dem dichotomen oder dem ordinalen Rasch-Modell zu (siehe Abb. 76). Vor allem bei den komplexeren Beurteilungsaufgaben C (Bewegungsvergleich), D (Fehleranalyse) und E (synchrone Einschätzung aller Bewegungsmerkmale) erhöht sich über das ordinale Auswertungsverfahren der durchschnittliche Personenwert deutlich, weil viele Testpersonen, wenn sie nicht mit dem Hauptexpertenurteil übereinstimmen, so zumindest das Nebenexpertenurteil gewählt haben. Beim dichotomen Rasch-Modell wird hingegen nicht unterschieden, ob die Testperson nah oder fern dem Hauptexpertenurteil liegt. Trotzdem bleibt die Tendenz bestehen, dass der durchschnittliche Personeneigenschaftswert der Testpersonen mit zunehmender Aufgabenkomplexität sinkt und der Bewegungsvergleich am schwierigsten auszuführen ist. Somit kann der Versuch, die Aufgaben des Beurteilungstests so zu konzipieren, dass diese zunehmend schwieriger werden, nach der Testauswertung der Einzelteile als gelungen bezeichnet werden.

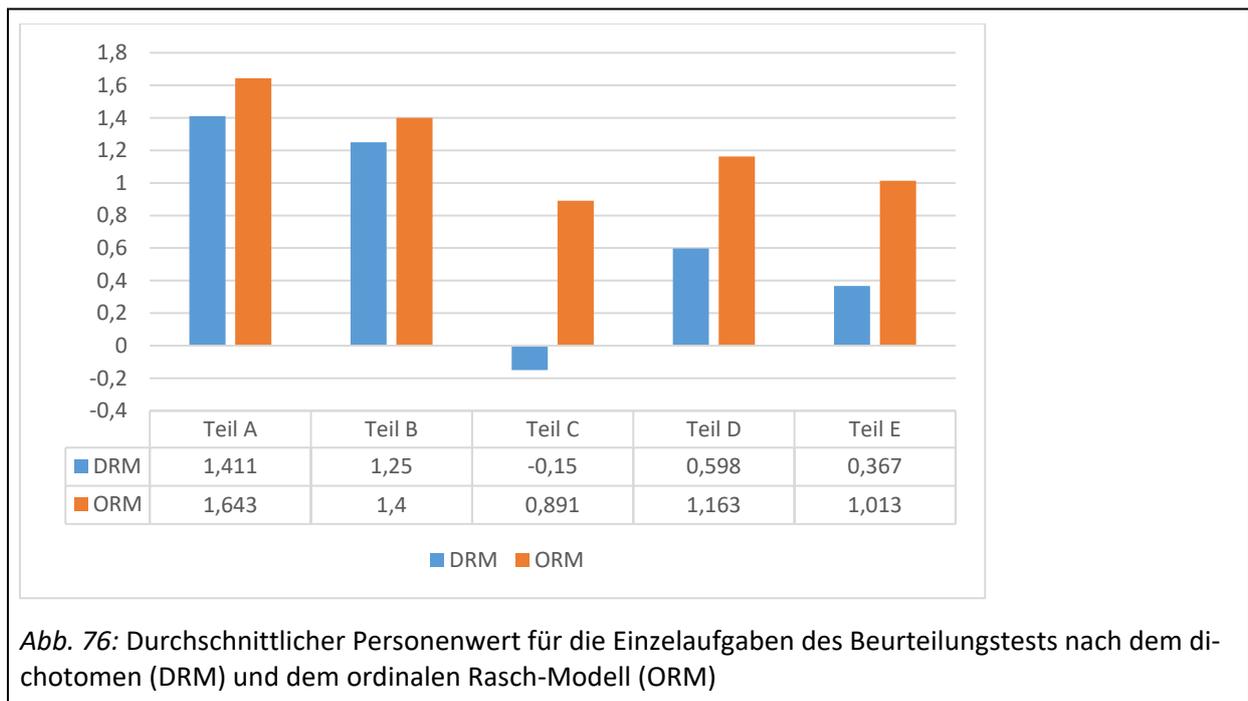


Abb. 76: Durchschnittlicher Personenwert für die Einzelaufgaben des Beurteilungstests nach dem dichotomen (DRM) und dem ordinalen Rasch-Modell (ORM)

Betrachtet man die lineare Korrelation zwischen den ermittelten Personenwerten nach dem dichotomen und dem ordinalen Rasch-Modell, so sieht man, dass die Übereinstimmung der Rangreihenfolge in den einzelnen Teilen relativ hoch ist, wobei bei Teil B die Auswirkungen der Analyse nach dem ORM-Modell am gravierendsten sind (siehe Tab. 32). In der Folge werden beide Personenwerte für die Analysen herangezogen, weil nicht eindeutig zu entscheiden ist, ob die Möglichkeit, die Annäherung an das Hauptexpertenurteil auch im Personenwert auszudrücken, die Präferenzierung des ordinalen Rasch-Modells rechtfertigt.

Tab. 32: Korrelationskoeffizienten zwischen der Auswertung nach dem dichotomen und dem ordinalen Rasch-Modell

BT_A_6	BT_B	BT_C	BT_D	BT_E	
B_A_ORM	BT_B_ORM	BT_C_ORM	BT_D_ORM	BT_E_ORM	
r = 0,861	r = 0,614	r = 0,794	r = 0,968	r = 0,780	

Überprüft man die Verteilung der Personenfähigkeitswerte der einzelnen Teile des Beurteilungstests (ausgewertet nach dem Partial-Credit-Modell) mit dem Shapiro-Wilk-Test, welcher besonders für kleine Stichproben (unter 100 Testpersonen) geeignet ist, so kommt man nach Ansicht der p-Werte zu dem Ergebnis, dass die Annahme einer Normalverteilung für alle Testteile verworfen werden muss. Der Kolmogorov-Smirnov-Test mit Signifikanzkorrektur nach Lilliefors bestätigt dieses Ergebnis (siehe Appendix, Tab. 62). Analysiert man die Daten zusätzlich mithilfe der graphischen Möglichkeiten (Histogramm mit Normalverteilungskurve, Q-Q-Diagramme und Boxplot), so ist erkennbar, dass die Normalverteilung für die Teile C, D und E aufgrund zu großer Häufigkeiten bei mittleren Personenparametern abgelehnt wird. Teile A und B haben speziell im Bereich des Umkehrpunkts der idealtypischen Normalverteilungskurve Lücken (siehe Abb. 77), während die Teile C, D und E mit je zwei Overachievern (siehe Abb. 78) mehrere Extremwerte aufweisen. Aufgrund dieser Erkenntnisse wird die weitere

Auswertung der Daten auf nicht-parametrische Weise erfolgen, wenngleich an geeigneter Stelle parallel auch parametrische Berechnungen eingesetzt werden, um die nicht-parametrischen Ergebnisse zusätzlich abzusichern oder für den Fall einer Normalverteilung zu widerlegen.

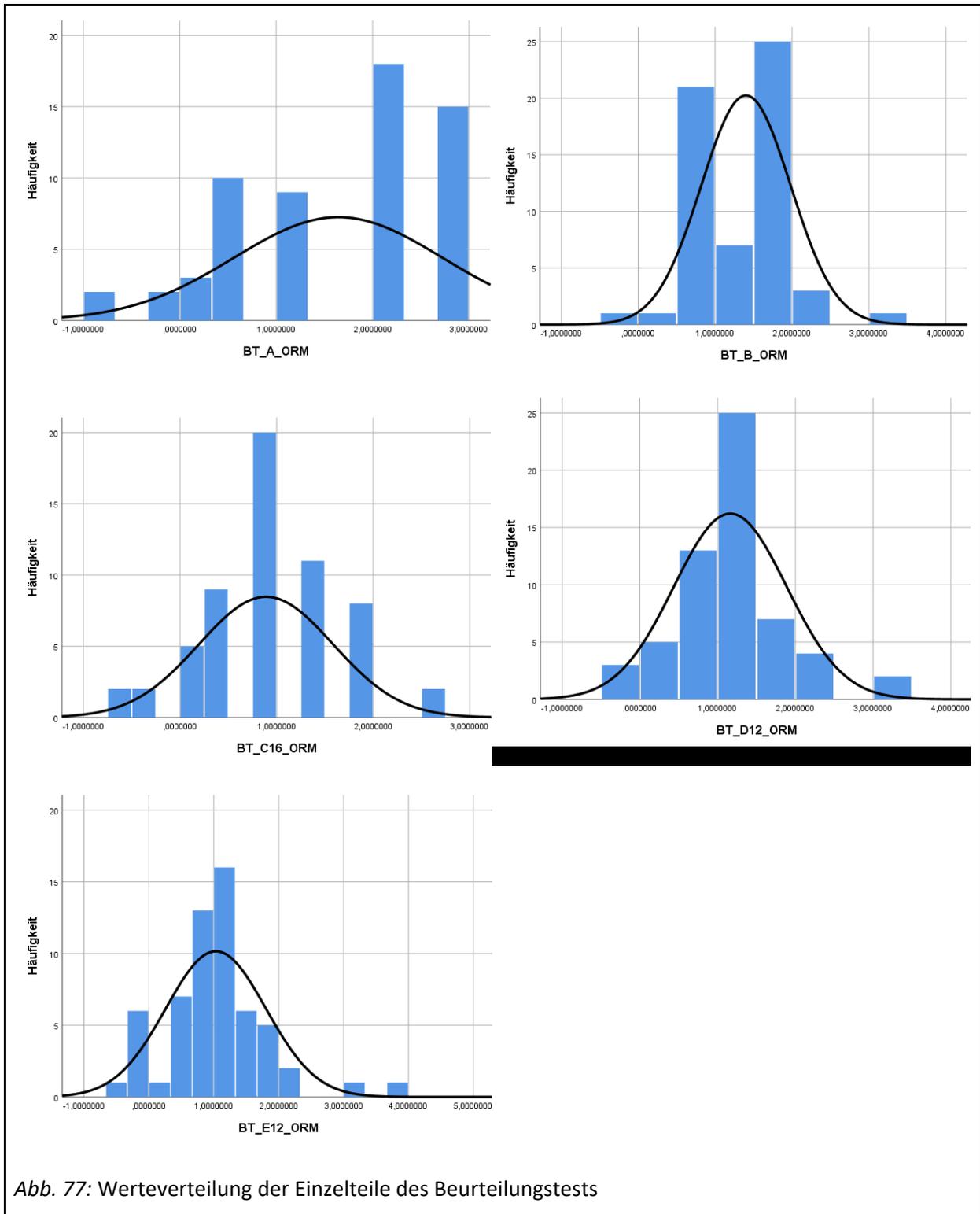


Abb. 77: Werteverteilung der Einzelteile des Beurteilungstests

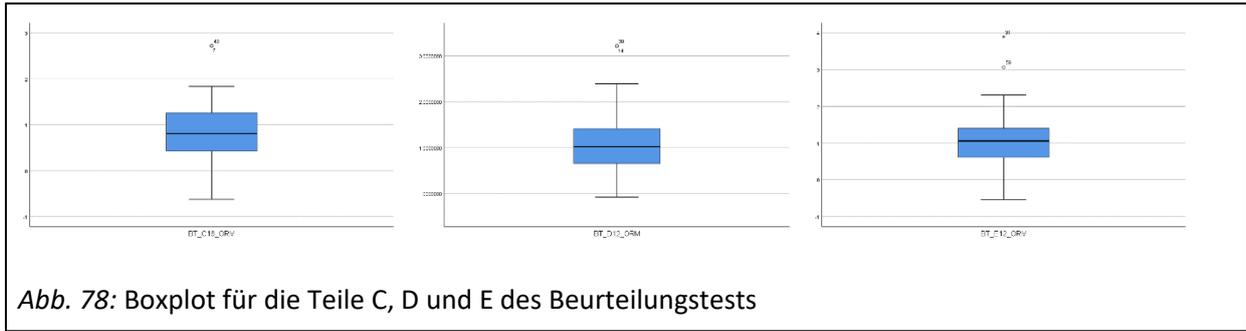


Abb. 78: Boxplot für die Teile C, D und E des Beurteilungstests

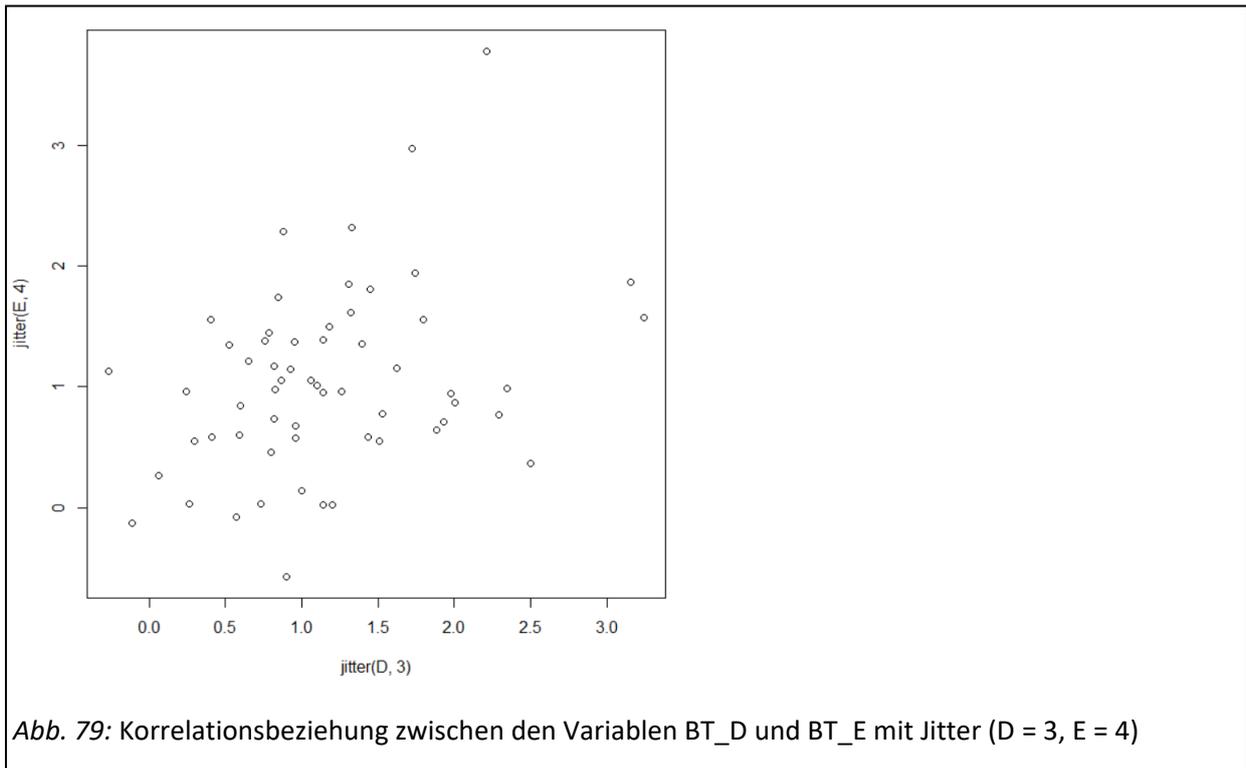
Interessant ist auch die Feststellung, dass zwischen den Personenwerten der komplexeren Teile C, D und E schwach positive lineare Zusammenhänge nachgewiesen werden konnten, während Teil A mit keinem anderen und Teil B lediglich mit Teil D überraschenderweise schwach negativ korreliert (siehe Tab. 33). Die stärkste Korrelation, welche auch in Abb. 79 graphisch anhand der Verteilung der Einzelwerte zu erkennen ist, zeigen die Teile D und E (mit $r = 0,342$, $p = 0,008$), welche sich sogar als Pearson's Produkt-Moment-Korrelation auf $r = 0,426$ erhöht. Dieses Ergebnis deutet darauf hin, dass für die Teile A (reine Zuordnung der Qualitätsstufen zu drei Bewegungsbeispielen) und B (Bestimmung der Qualitätsstufe nur eines Bewegungsmerkmals) andere Beurteilungsqualitäten verlangt sind als bei den komplexeren Teilen C (Bewegungsvergleich), D (Fehleranalyse) und E (Gesamtanalyse). Ähnlich der Unterscheidung des TKBV verlangen die Teile A und B ein Wiedererkennen eines einzelnen Bewegungsmerkmals, während die Teile C, D und E das freie Beurteilen, Vergleichen und Analysieren teils mehrerer Bewegungsmerkmale erfordert. Daher ist es denkbar, für die Teile C, D und E (Variable BT_CDE) einen gemeinsamen Personenparameter zu berechnen und diesen bei weiteren Zusammenhangsanalysen als Gesamtvariable zusätzlich zu berücksichtigen.

Tab. 33: Nichtparametrische Korrelationen zwischen den nach dem dichotomen Rasch-Modell geschätzten Personenwerten der Einzelteile des Beurteilungstests (nach Spearman)

Nach Spearman		B	C	D	E
A	r	0,039	0,009	0,027	-0,084
	Sig. (2-seitig)	0,769	0,948	0,838	0,529
	N	59	59	59	59
B	r		-0,014	-0,226	-0,054
	Sig. (2-seitig)		0,914	0,086	0,687
	N		59	59	59
C	r			0,238	,258*
	Sig. (2-seitig)			0,069	0,048
	N			59	59
D	r				,342**
	Sig. (2-seitig)				0,008
	N				59

*. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig).

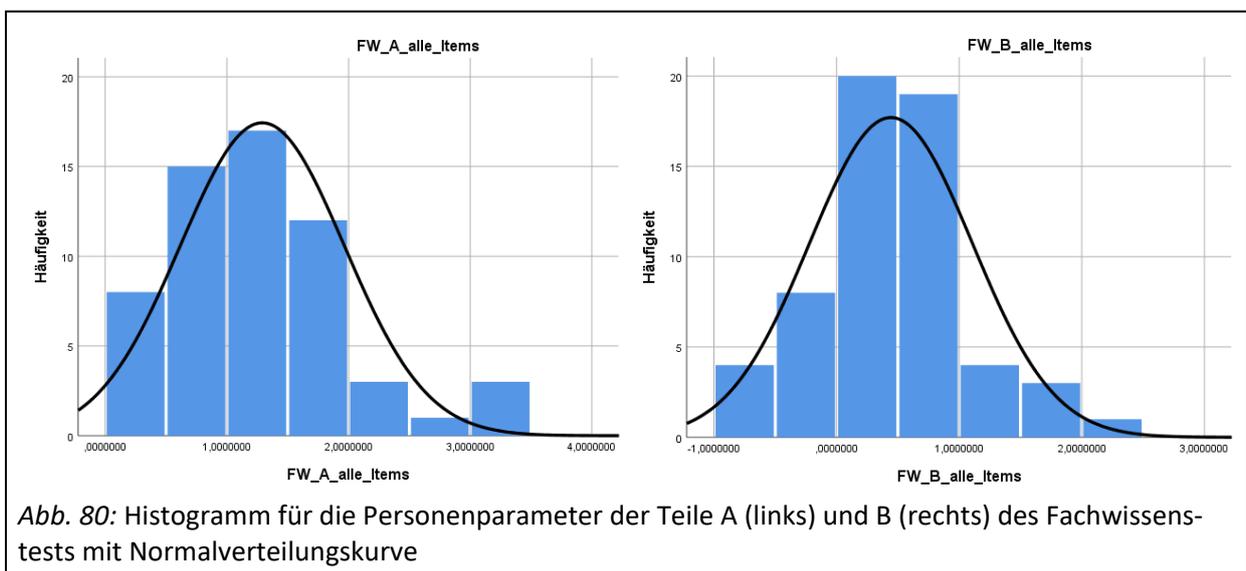
** . Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig).



5.1.2 Fachwissen

5.1.2.1 Zusammenfassung der Testergebnisse

Die Auswertung des Fachwissenstests (siehe Kapitel 5.2.1) zeigt, dass die Testpersonen über den räumlich-zeitlichen Bewegungsverlauf (Teil A) besser Bescheid wissen als über biomechanische Zusammenhänge (Teil B). Bei Teil A wurde mit $\theta_A = 1,291$ (SD = 0,675) ein deutlich höherer durchschnittlicher Personenparameterwert erzielt als bei Teil B mit $\theta_B = 0,443$ (SD = 0,665). Die graphisch dargestellte Häufigkeit der Messwerte zeigt in Abb. 80 für beide Testteile eine Annäherung an die Normalverteilung, wobei diese rechnerisch nach dem Shapiro-Wilk-Test für Teil A mit $p = 0,006$ abgelehnt und für Teil B mit $p = 0,201$ angenommen werden darf.



Die Auswertung des Fehlerwissens erfolgt auf deskriptive Weise, da aufgrund des freien Antwortformats keine Auswertung mit dem Rasch-Modell möglich ist. Die Testpersonen erzielten durchschnittlich 6,07 Punkte (SD = 2,00) beim Bewegungsabschnitt Anlauf (entspricht 60,7 Prozent von zehn möglichen Punkten), 7,20 Punkten (SD = 2,22) beim Abwurf (entspricht 72,0 Prozent von zehn möglichen Punkten) und 2,36 Punkte (SD = 0,99) beim Abfangen (entspricht 47,2 Prozent von fünf möglichen Punkten). Somit ist das Fehlerwissen über die Hauptphase der Wurfbewegung etwas ausgeprägter als über die Vorbeschleunigungsphase. Das Fehlerwissen über die Endphase des Abfangens ist hingegen deutlich geringer ausgeprägt. Insgesamt erreichen die Testpersonen bei Zusammenfassung der drei Teile 15,63 (SD = 3,76) von maximal 25 erreichbaren Punkten, wobei der Spitzenwert von 23 Punkten drei Mal, der Tiefstwert von acht Punkten zwei Mal erreicht wurde (siehe Abb. 81).

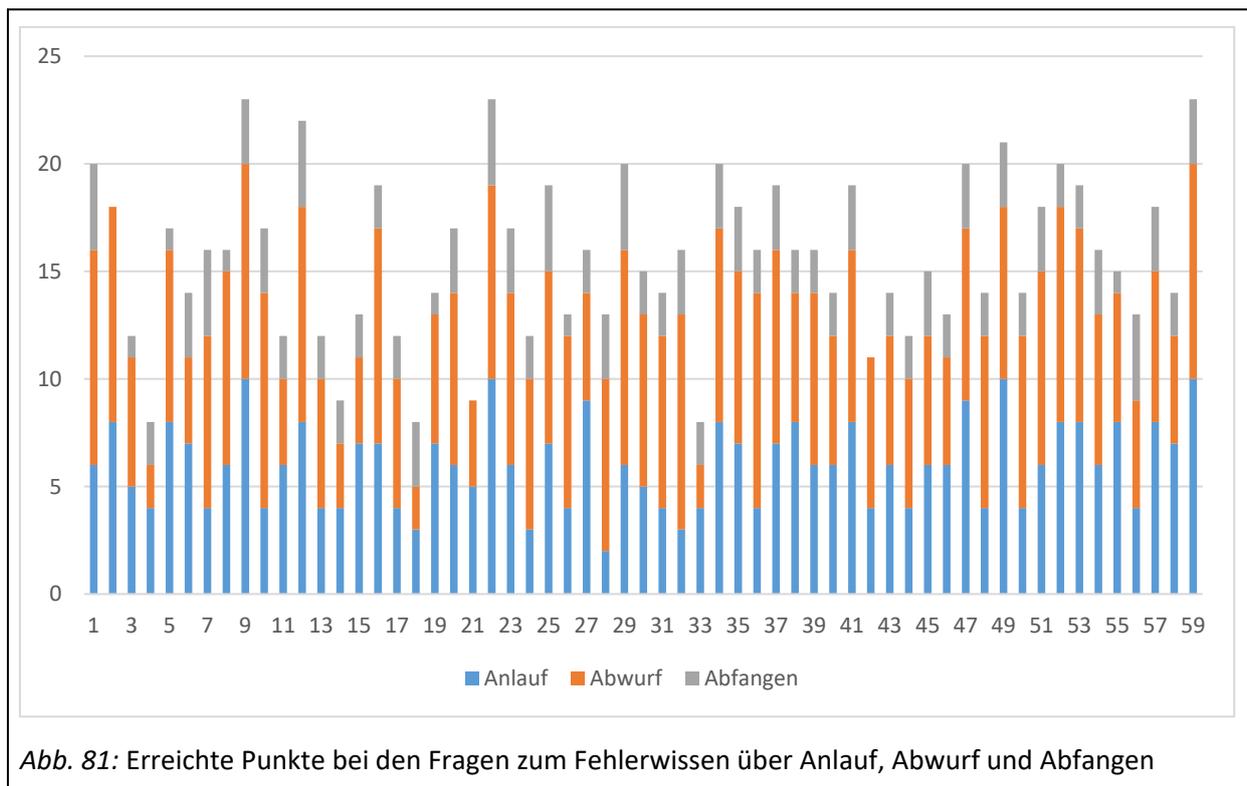


Abb. 81: Erreichte Punkte bei den Fragen zum Fehlerwissen über Anlauf, Abwurf und Abfangen

Die erreichten Punktzahlen korrelieren dabei zwischen Anlauf und Abwurf mit $r = 0,270$ ($p = 0,039$), zwischen Abwurf und Abfangen mit $r = 0,319$ ($p = 0,014$) und zwischen Anlauf und Abfangen mit $r = 0,146$ ($p = 0,271$). Es zeigt sich also, dass das Fehlerwissen über die Hauptfunktionsphasen mit den beiden Nebenfunktionsphasen deutlich stärker zusammenhängt als das Fehlerwissen zwischen den beiden Nebenfunktionsphasen. Die Testpersonen, welche viel über Bewegungsabweichungen in der Hauptphase wissen, kennen tendenziell auch mehr Fehlerbilder in der Vorbereitungs- und Endphase der Wurfbewegung.

Die Auswertung der Bildersortieraufgabe zur Feststellung der visuellen Bewegungsvorstellung ergibt eine durchschnittlich erreichte Punktzahl von 6,86 (SD = 1,84) (bei maximal acht möglichen Punkten). Allein 35 Testpersonen erreichten den Maximalwert und lediglich acht Testpersonen haben vier oder weniger Punkte. Diese Ergebnisse sprechen dafür, dass der Testgruppe die Aufgabenstellung der Bildersortierung insgesamt leichtfällt und ein sehr ausgeprägtes und stabiles Wissen über die Reihenfolge der ausgewählten Schlüsselpositionen vorhanden ist.

5.1.2.2 *Zusammenhang zwischen Fachwissen und Beurteilungsleistung*

Überprüft man die Korrelation zwischen den Einzelvariablen des Fachwissens und den einzelnen Teilen des Beurteilungstests, so sind die meisten Beziehungen nur mit einem sehr schwachen Korrelationskoeffizienten zu beschreiben. Das räumlich-zeitliche Bewegungswissen (Teil A des Fachwissenstests) übt unabhängig von den oben genannten unterschiedlichen Auswertungsmöglichkeiten auf keine Beurteilungsaufgabe einen nennenswerten Einfluss aus. Einzig mit Teil E des Beurteilungstests besteht ein leicht negativer Zusammenhang ($r = -0,113$ mit $p = 0,392$), d. h. Personen mit einem geringeren räumlich-zeitlichen Bewegungswissen haben bei der synchronen Gesamtanalyse der Bewegungsmerkmale in dieser Studie besser abgeschnitten als die Personen mit einem höheren RZ-Wissen. Bezüglich des biomechanischen Wissens lassen sich etwas zuverlässigere Aussagen treffen. Mit den Beurteilungsaufgaben B, C und D gibt es zwar keinen nachweislichen Zusammenhang, aber mit Teil A besteht eine positive lineare Korrelation ($r = 0,250$ mit $p = 0,057$). Das biomechanische Wissen scheint den Testpersonen die richtige Zuordnung der Bewegungsbeispiele zu den Qualitätsstufen zu erleichtern. Der signifikant negative Zusammenhang zwischen dem biomechanischen Wissen und dem Teil E ($r = -0,217$ mit $p = 0,098$) bestätigt die eben genannte Vermutung, dass die Lösung der komplexen Beurteilungsaufgabe der synchronen Gesamtanalyse nicht auf einem ausgeprägten biomechanischen oder räumlich-zeitlichen Wissen beruht.

Das Fehlerwissen korreliert positiv mit Teil A ($r = 0,129$ mit $p = 0,332$), Teil B ($r = 0,171$ mit $p = 0,194$) und Teil D ($r = 0,168$ mit $p = 0,205$), wenngleich für den Zusammenhang keine Signifikanz nachgewiesen werden konnte. Das Wissen über vom Technikleitbild abweichende Bewegungsausführungen scheint aber für die drei verschiedenen Beurteilungsaufgaben ein zumindest gleichbedeutender, wenn nicht sogar gewinnbringenderer Einflussfaktor zu sein als das reine Fachwissen. Dies erscheint einleuchtend, weil in den einzelnen Qualitätsstufen die Bewegungsabweichungen eingearbeitet sind und ein gewisses Vorwissen helfen kann, die entsprechenden Fehlerbilder dann auch bei schnell ablaufenden Bewegungen zu erkennen und damit das eigene Beurteilungsergebnis zu unterstützen.

Die konkrete Bewegungsvorstellungsfähigkeit hingegen, Bilder von Schlüsselpositionen der Wurfbewegung in die richtige Reihenfolge bringen zu können, korreliert mit den Beurteilungsaufgaben A ($r = 0,085$ mit $p = 0,524$) und B ($r = 0,173$ mit $p = 0,189$) schwach positiv, mit den komplexeren Beurteilungsaufgaben C ($r = -0,206$ mit $p = 0,118$), D ($r = -0,153$ mit $p = 0,246$) und E ($r = -0,207$ mit $p = 0,116$) jedoch konstant negativ, wenngleich keine Signifikanz behauptet werden kann. Dies ist nicht unbedingt überraschend, weil die Kenntnis von Keypoints und deren Reihung im Bewegungsverlauf sehr eng mit dem räumlich-zeitlichen Bewegungswissen verknüpft ist, welches wie oben erwähnt beispielsweise mit Teil E auch schwach negativ korreliert. Dieser statisch-bildhafte und sehr kognitive Zugang zur Schlagwurfbewegung hilft den Testpersonen offensichtlich bei den einfachen Beurteilungsaufgaben, die aufgrund ihrer Beschränkung auf ein zu beobachtendes Bewegungsmerkmal möglicherweise noch recht bewusstseinsfähig geschehen kann. Überraschend ist es aber trotzdem, dass diese visuelle Bewegungsvorstellungsfähigkeit dann kontraproduktiv zu sein scheint für die Beurteilungsaufgaben, bei denen Informationsaufnahme und -verarbeitung des Beurteilungsprozesses wesentlich komplexer ablaufen müssen. Fehlerwissen und Bildreihe selbst hingegen weisen keine Beziehung zueinander auf ($r = 0,07$ mit $p = 0,959$).

Diese Zusammenhänge stützen die Vermutung, dass die Komplexität der Beurteilungsaufgabe ein entscheidender Faktor ist. Je einfacher die Beurteilungsaufgabe ist, desto eher kann das Wissen über Fehlerbilder sowie das Wissen über die Reihenfolge visuell präsentierter Schlüsselpositionen die Einschätzung der Bewegungsqualität befruchten. Bei komplexeren Beurteilungsaufgaben spielen diese Wissensbereiche dann wieder keine Rolle, so dass sogar Personen mit einer schlechteren visuellen Bewegungsvorstellung eine höhere Lösungswahrscheinlichkeit der Beurteilungsaufgaben erreichen.

5.1.3 Fähigkeit zum Ist-Sollwert-Vergleich

5.1.3.1 Zusammenfassung der Testergebnisse

Wenn man die beiden Teile des TKBV gemeinsam auswertet, so ergibt sich ein durchschnittlicher Personenparameterwert von 1,400 (SD = 1,042). Dieser hohe Personenfähigkeitswert lässt die Annahme zu, dass die Anforderungen des TKBV von den Testpersonen sehr gut gemeistert werden konnten. Anhand der graphischen Häufigkeitsdarstellung in Abb. 82 lässt sich erkennen, dass zwei Messwerte die Normalverteilungskurve deutlich überschreiten. Trotzdem bestätigt der Shapiro-Wilk-Test die Annahme, dass die Normalverteilung für den TKBV in dieser Studie angenommen werden darf ($p = 0,144$).

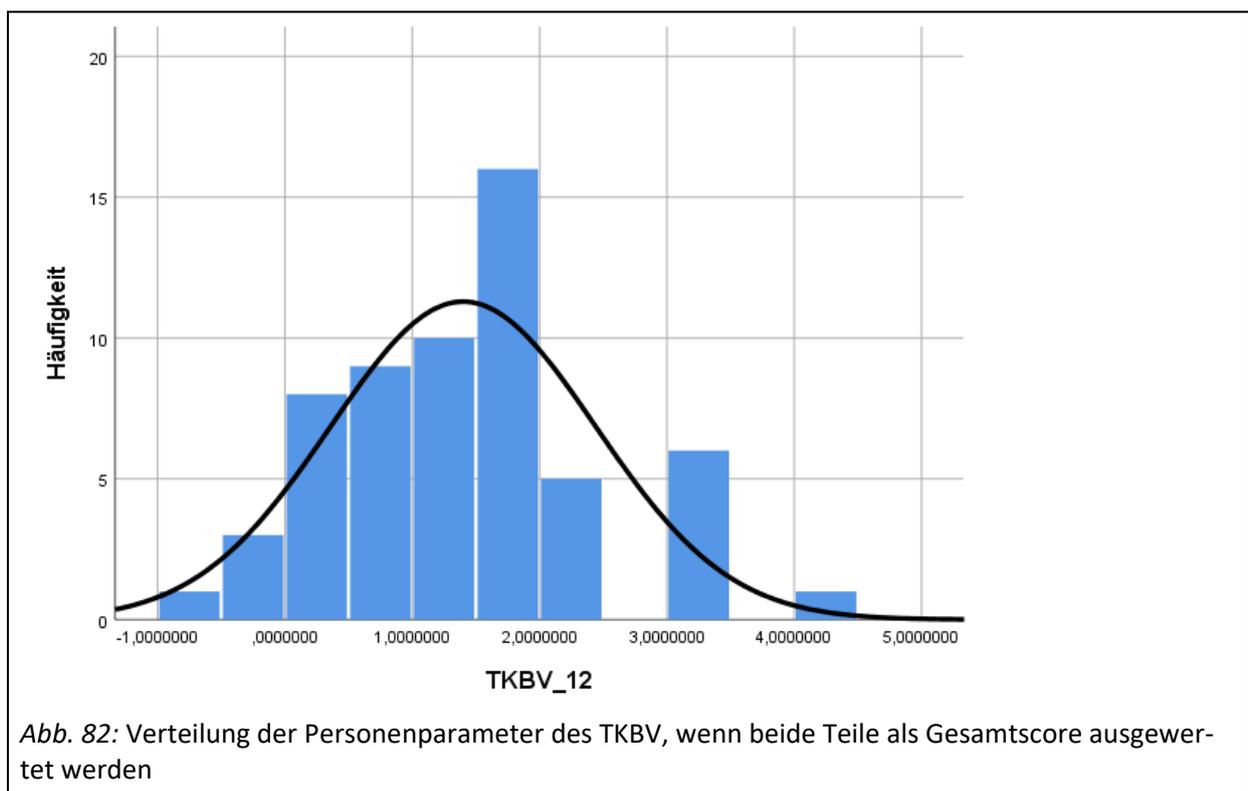


Abb. 82: Verteilung der Personenparameter des TKBV, wenn beide Teile als Gesamtscore ausgewertet werden

Wenn man die beiden Teile des TKBV einzeln auswertet, so ergibt sich ein differenzierteres Bild der Personenfähigkeiten. Bei Teil 1 erreichen die Testpersonen mit $\theta = 1,974$ (SD = 1,053) einen deutlich höheren durchschnittlichen Personenparameter als bei Teil 2 mit $\theta = 0,839$ (SD = 1,256). Da die graphische Analyse der Häufigkeitsverteilung noch keine Rückschlüsse auf die Normalverteilung zulässt (siehe Abb. 83), wird mit dem Shapiro-Wilk-Test berechnet, dass für Teil 1 eine Normalverteilung abgelehnt werden muss ($p = 0,000$) und für Teil 2 eine solche angenommen werden darf ($p = 0,165$). Insgesamt können die Wiedererkennungsaufgaben von Teil 1 sehr zuverlässig gelöst werden, während die

freie Wiedergabe der Körperendpositionen bei Teil 2 wesentlich mehr Testpersonen bereits überfordert. Zwischen den erreichten Personenparametern der Teile 1 und 2 besteht eine schwach positive Korrelation ($r = 0,211$ mit $p = 0,109$), eine lineare Abhängigkeit zwischen den beiden Teilen kann jedoch nicht nachgewiesen werden, weil das Regressionsmodell insgesamt nicht signifikant ist ($F(1;57) = 3,318$, $p = 0,074$).

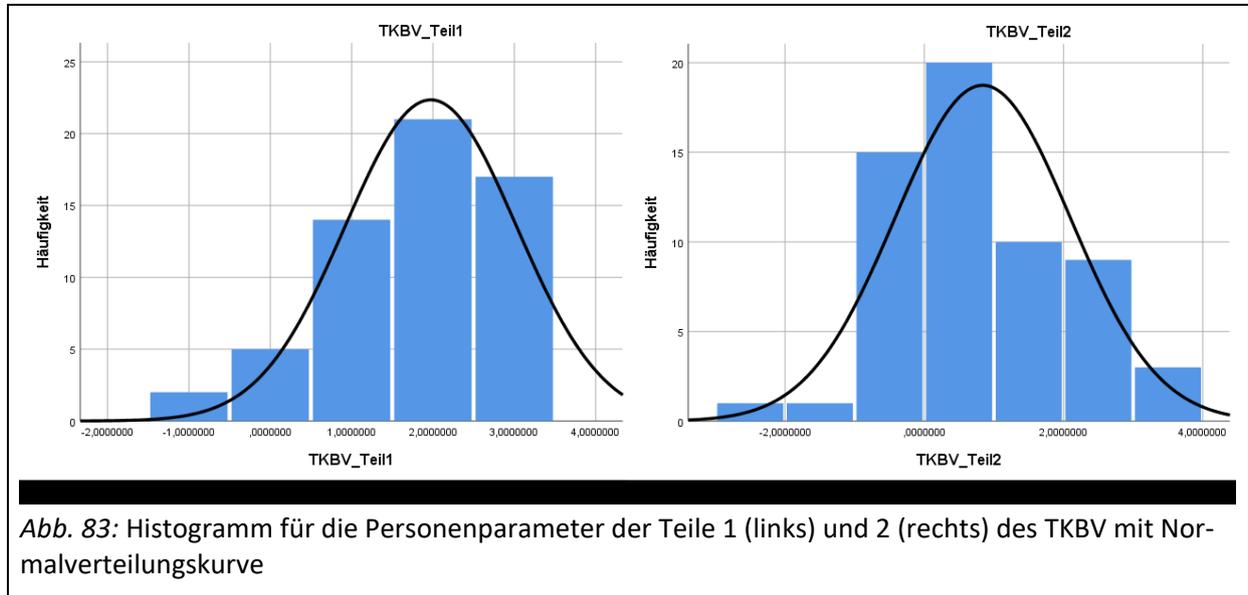


Abb. 83: Histogramm für die Personenparameter der Teile 1 (links) und 2 (rechts) des TKBV mit Normalverteilungskurve

5.1.3.2 Zusammenhang zwischen TKBV und Beurteilungsleistung

Setzt man die Personenwerte des TKBV in Beziehung zu den Personenwerten des Beurteilungstests, so erkennt man einen ähnlichen Trend wie bei den Variablen Fachwissen, Fehlerwissen und visuelle Bewegungsvorstellungsfähigkeit. Mit dem Gesamtpersonenwert beider Teile des TKBV (Variable TKBV_12) können keine eindeutigen Zusammenhänge mit den Beurteilungsaufgaben nachgewiesen werden. Lediglich mit Teil A des Beurteilungstests besteht eine leicht positive Korrelation ($r = 0,148$ mit $p = 0,262$). Der TKBV kann offensichtlich ein schwacher Indikator für das Lösen der Beurteilungsaufgabe von Teil A sein, Bewegungsausführungen einer Qualitätsstufe eines Bewegungsmerkmals zuzuordnen. Bei komplexeren Beurteilungsaufgaben ist diese Indikatorfunktion nicht mehr gegeben und kehrt sich sogar ins Gegenteil um, so dass ein guter TKBV-Wert sogar auf eine schlechtere Beurteilungsleistung hinweisen kann.

Wie eben erwähnt, kann zwischen den erreichten Personenparametern der TKBV-Teile 1 und 2 auch nur ein Korrelationskoeffizient von $r = 0,211$ ($p = 0,109$) festgestellt werden, so dass eine getrennte Betrachtung der beiden Teile sinnvoll erscheint. Hierbei fällt vor allem die negative Beziehung zwischen Teil 1 des TKBV und dem Beurteilungskomplex C ($r = -0,196$ mit $p = 0,137$), D ($r = -0,192$ mit $p = 0,146$) und E ($r = -0,144$ mit $p = 0,277$) auf. Testpersonen, welche die Wiedererkennungsaufgaben von Teil 1 des TKBV (das Bestimmen der richtigen Körperendposition nach gedanklichem Nachvollziehen bestimmter Bewegungsabfolgen) gut bewältigen, erzielen bei komplexeren Beurteilungsaufgaben wie dem Bewegungsvergleich oder der Fehleranalyse tendenziell schlechtere Ergebnisse als Testpersonen mit geringeren Personenwerten des TKBV-Teil 1. Die Personenparameter von Teil 2 des TKBV spiegeln diese Tendenz nicht wieder, sondern ergeben leicht positive oder zumindest neutrale r -Werte mit allen Teilen des Beurteilungstests, wobei vor allem Teil A ($r = 0,169$ mit $p = 0,202$) und Teil C (mit $r = 0,239$ mit $p = 0,068$) hervorzuheben sind (siehe Tab. 34).

Tab. 34: Korrelationen zwischen den Teilen des Beurteilungstests (ausgewertet nach dem ORM) und den Teilen des TKBV (ausgewertet nach dem DRM)

		BT_A ORM	BT_B ORM	BT_C ORM	BT_D ORM	BT_E ORM
TKBV_12	r	0,165	0,024	0,089	-0,069	-0,028
	Sig. (2-seitig)	0,213	0,858	0,501	0,602	0,833
	N	59	59	59	59	59
TKBV_Teil1	r	0,045	-0,019	-0,196	-0,192	-0,144
	Sig. (2-seitig)	0,738	0,885	0,137	0,146	0,277
	N	59	59	59	59	59
TKBV_Teil2	r	0,169	0,036	0,239	0,020	0,023
	Sig. (2-seitig)	0,202	0,786	0,068	0,881	0,861
	N	59	59	59	59	59

Insgesamt lässt sich festhalten, dass die Personenwerte, die mit dem TKBV erhoben wurden, keinen signifikanten Zusammenhang zu den Personenwerten der einzelnen Teile des Beurteilungstests aufweisen. Die Fähigkeit des Wiedererkennens vorgestellter Bewegungsmerkmale (TKBV Teil 1) weist sogar eine schwach negative Korrelation zu den komplexeren Beurteilungsteilen (C, D und E) auf. Die Fähigkeit zur freien Wiedergabe von Bewegungsvorstellungen (TKBV Teil 2) korreliert lediglich mit den Teilen A und C des Beurteilungstests schwach positiv. Diese Feststellungen sind durchaus überraschend, denn die Fähigkeit, sich eine Bewegung mental vorzustellen und dann in unterschiedlicher Form abzurufen, weist eine (wenngleich sehr vereinfachte) Nähe zum Prozess der Bewegungsbeurteilung auf. Zumindest scheint diese vermutete Basisfähigkeit nicht über den Test zur Kontrolle der Bewegungsvorstellungsfähigkeit abbildbar zu sein. Der hohe Korrelationskoeffizient zwischen Teil 2 des TKBV und Aufgabe C des Beurteilungstests kann möglicherweise auch über eine erhöhte Merkfähigkeit der Testpersonen erklärt werden.

Betrachtet man die Ergebnisse des Bewegungsbeobachtungstests, so fallen zwei schwach positive Korrelationen mit den Beurteilungsaufgaben C ($r = 0,145$ mit $p = 0,272$) und E ($r = 0,110$ mit $p = 0,408$) auf, während die anderen Teile des Beurteilungstests keine nennenswerten Korrelationen aufweisen. Es bestätigt sich immerhin die leichte Tendenz, dass ein guter Personenparameter beim Bewegungsbeobachtungstest grundsätzlich positiv mit den komplexeren Beurteilungsaufgaben korreliert. Die Aufgaben des Bewegungsbeobachtungstests, bei dem zwei Bewegungsausführungen miteinander verglichen und Abweichungen voneinander erkannt werden müssen, weisen eine inhaltliche Nähe zu Teil C des Beurteilungstests auf, in dem es ebenfalls um den Vergleich der qualitativen Ausführung von Bewegungsmerkmalen geht. Es kann also durchaus begründet vermutet werden, dass die Fähigkeiten, die für das Vergleichen von Bewegungsmerkmalen erforderlich sind, mit Hilfe von sehr stark abstrahierten Übungsaufgaben wie denen des Bewegungsbeobachtungstests verbessert werden können.

Überraschender Weise korrelieren die Personenparameter des Bewegungsbeobachtungstests auf dem 0,01-Niveau signifikant mit den Personenparametern von TKBV-Teil 1 ($r = 0,343$ mit $p = 0,008$) und TKBV-Teil 2 ($r = 0,340$ mit $p = 0,008$), mit dem Gesamtergebnis des TKBV (TKBV_12) sogar mit $r = 0,406$ ($p = 0,001$). Die Fähigkeiten, die für eine gute Bearbeitung der beiden Teile des TKBV nötig sind, werden demnach auch bei den vergleichenden Aufgaben des Bewegungsbeobachtungstests verlangt. Somit bestätigt sich die Umsetzung, den Bewegungsbeobachtungstest in enger Anlehnung an den TKBV zu konstruieren und ähnliche Fähigkeiten in einer komplexeren Aufgabenstellung, die durch bewegte Bilder sowie gelegentliche Perspektivwechsel zu erreichen versucht wurde, anzusprechen.

5.1.4 Beurteilungserfahrung

5.1.4.1 Zusammenfassung der Testergebnisse

Bei der Auswertung des Fragebogens über die Vorerfahrung in der Beurteilung sportlicher Bewegungen handelt es sich um dichotome Daten. Die Testpersonen mussten mit ja oder nein antworten, ob und in welcher Form sie schon sportliche Bewegungen beurteilt haben (Fragebogen siehe Abb. 119 im Appendix). Eine rein deskriptive Auswertung der Daten ergibt, dass von den 59 Probandinnen und Probanden neun bereits Beurteilungserfahrungen mit leichtathletischen Wurfbewegungen gemacht haben. 31 Personen (darunter sechs, die auch schon leichtathletische Beurteilungserfahrung haben) geben an, schon Schlag(wurf)bewegungen in anderen Sportarten wie Tennis, Handball oder Volleyball beurteilt zu haben, während 32 Probandinnen und Probanden (darunter 19, die auch schon andere Schlag(wurf)bewegungen bewertet haben) schon Bewegungsbeurteilungen in Sportarten durchgeführt haben, die keinen Bezug zu Schlag(wurf)bewegungen haben wie Turnen, Fußball oder Basketball. Es bleibt eine Gruppe von 14 Personen übrig, die noch überhaupt keine Beurteilungserfahrung sportlicher Bewegungen gemacht hat. Weitere Gruppendifferenzierungen wären möglich, können jedoch in dieser Studie aufgrund einer zu kleinen sich ergebenden Gruppengröße nicht analysiert werden. Dies liegt vorwiegend an der geringen Probandenzahl, die einen Bezug zur Beurteilung leichtathletischer Bewegungen hat. Aus diesem Grund bleibt auch eine Testperson, die nur in der Leichtathletik und sonst keiner anderen Sportart Beurteilungserfahrungen hat, in den weiteren Analysen bezüglich der Beurteilungserfahrung unberücksichtigt, weil sie keiner der vorliegenden Gruppen eindeutig zugeordnet werden konnte und für eine eigene Gruppe, die in dieser Studie zwischen Stufe 4 und Stufe 5 liegen würde, keine weiteren Testpersonen in Frage kommen. Somit können folgende Stufen der Vorerfahrung anhand der erhobenen Daten unterschieden werden (siehe auch Abb. 84):

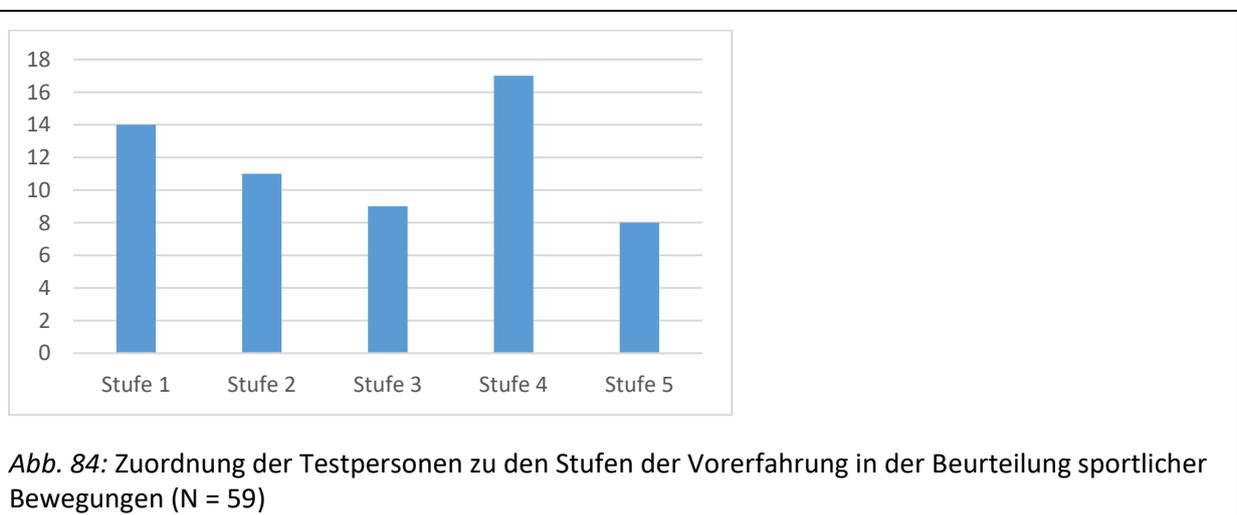
Stufe 1: keinerlei Vorerfahrung im Beurteilen sportlicher Bewegungen (N = 14)

Stufe 2: Vorerfahrungen in Sportarten, in denen keine Schlag(wurf)bewegungen vorkommen (N = 11)

Stufe 3: Vorerfahrungen in Sportarten, in denen Schlag(wurf)bewegungen vorkommen (N = 9)

Stufe 4: Vorerfahrungen in Sportarten, in denen Schlag(wurf)bewegungen vorkommen, und Vorerfahrungen in Sportarten, in denen keine Schlag(wurf)bewegungen vorkommen (N = 17)

Stufe 5: Vorerfahrungen in der Beurteilung leichtathletischer Schlagwurfbewegungen und Vorerfahrungen in anderen Sportarten (N = 8)



5.1.4.2 Zusammenhang zwischen der Beurteilungserfahrung und der Beurteilungsleistung

Um den Einfluss der Beurteilungserfahrung auf die in dieser Studie gemessene Beurteilungsleistung zu untersuchen, wurde ein nicht-parametrischer Gruppenvergleich durchgeführt. Bevor dieser Gruppenvergleich auch statistisch überprüft wird, werden die Personenmittelwerte der einzelnen Teile des Beurteilungstests für die einzelnen Stufen der Beurteilungsvorerfahrung gebildet. Für eine schnellere Übersicht wurden die Felder der Tab. 35 grün gefärbt, die über dem Gesamtmittelwert liegen, wohingegen sich die roten Felder unter dem Gruppendurchschnitt befinden.

Tab. 35: Mittelwerte und Standardabweichung der Personenwerte der Beurteilungsaufgaben (berechnet nach dem ORM) für die einzelnen Stufen der Beurteilungsvorerfahrung

Vorerf. Gruppe		BT_A	BT_B	BT_C	BT_D	BT_E	BT_ABC DE	BT_AB	BT_CDE
Stufe 1	Mittelwert	1,448	1,436	0,908	1,405	1,139	1,135	1,343	1,071
	N	14	14	14	14	14	14	14	14
	SD	0,919	0,470	0,473	0,915	0,936	0,372	0,374	0,521
Stufe 2	Mittelwert	2,057	1,555	1,000	0,801	0,913	1,044	1,630	0,845
	N	11	11	11	11	11	11	11	11
	SD	1,125	0,691	0,811	0,546	0,681	0,300	0,730	0,404
Stufe 3	Mittelwert	1,587	1,137	0,609	0,998	0,821	0,864	1,184	0,757
	N	9	9	9	9	9	9	9	9
	SD	1,203	0,467	0,742	0,556	0,958	0,425	0,529	0,608
Stufe 4	Mittelwert	1,540	1,559	0,981	1,076	1,144	1,131	1,451	1,037
	N	17	17	17	17	17	17	17	17
	SD	1,256	0,638	0,744	0,610	0,565	0,313	0,640	0,411
Stufe 5	Mittelwert	1,697	1,084	0,835	1,605	1,005	1,053	1,181	1,056
	N	8	8	8	8	8	8	8	8
	SD	0,831	0,470	0,779	0,762	0,864	0,307	0,432	0,578
Insgesamt	Mittelwert	1,643	1,400	0,891	1,163	1,032	1,064	1,381	0,969
	N	59	59	59	59	59	59	59	59
	SD	1,082	0,581	0,695	0,726	0,773	0,344	0,568	0,492

Vergleicht man die Mittelwerte der Personeneigenschaften der einzelnen Stufen mit dem Gesamtdurchschnittswert über alle Gruppen hinweg, so kann man einige Auffälligkeiten formulieren. Die Gruppenmitglieder der Stufe 3 erzielten in allen Aufgabenteilen nur unterdurchschnittliche Werte und stellen somit die schwächste Teilgruppe dar. Die Testpersonen der Vorerfahrungsstufe 1 hingegen, die noch überhaupt keine Beurteilungserfahrungen sportlicher Bewegungen gesammelt haben, erreichen lediglich bei Teil A unterdurchschnittliche Ergebnisse, dafür wird die komplexe Aufgabe D überdurchschnittlich gut bewältigt. Abgesehen von einem schwachen Testteil B erreichen auch die beurteilungserfahrensten Probandinnen und Probanden von Stufe 5 nur im Testteil D auffällig überdurchschnittliche Mittelwerte. Testpersonen der Stufe 4 erzielen in den Aufgabenteilen B und E knapp den besten Durchschnittswert, insgesamt liegen alle Teile relativ nah am Gesamtdurchschnitt. Die Gruppe der

Stufe 2 zeigt einen idealtypischen Verlauf. Während die Aufgaben A (mit Abstand höchster Mittelwert der Gruppen), B und C noch überdurchschnittlich gelöst werden konnten, werden die komplexeren Teile D und E nicht mehr so erfolgreich bearbeitet.

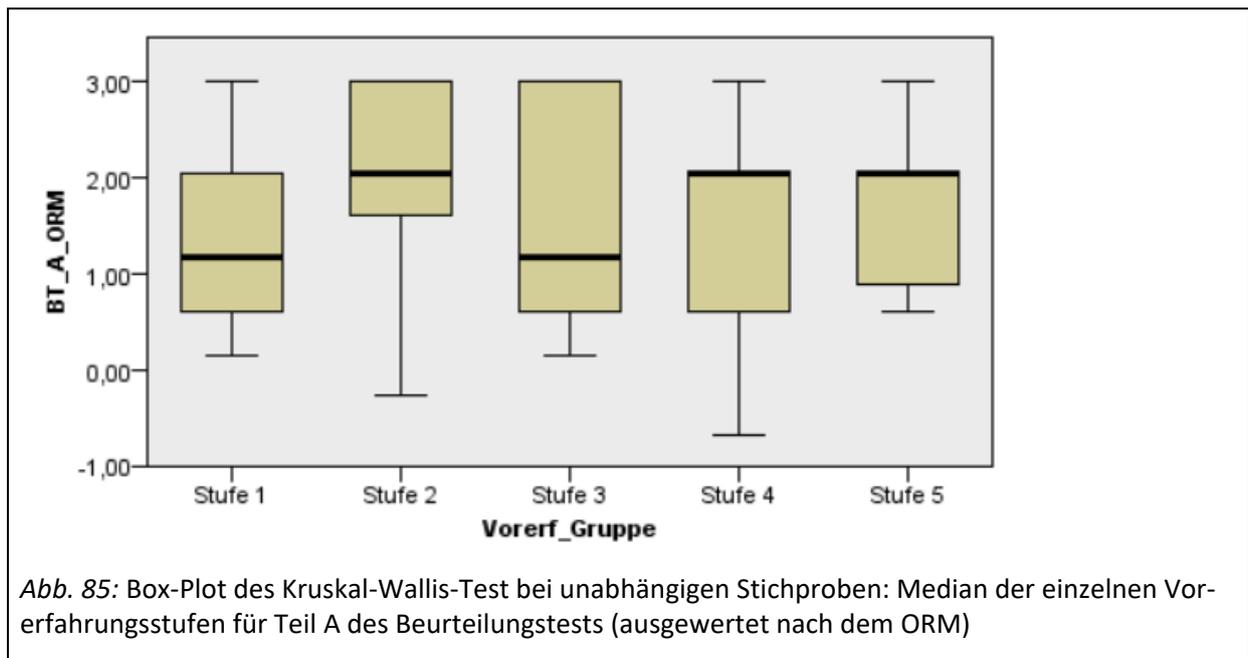
Führt man den Kruskal-Wallis-H-Test mit der Testvariable Personenwert des Gesamtbeurteilungstests (BT_ABCDE) durch, so ergibt sich ebenso ein nicht-signifikanter Wert ($p = 0,311$) wie für den gemeinsam berechneten Personenwert BT_AB ($p = 0,439$) sowie BT_CDE ($p = 0,275$). Die Nullhypothese, alle Stufen der Beurteilungsvorerfahrung seien hinsichtlich der erreichten Personenwerte des Beurteilungstests gleich, ist somit gültig. Überprüft man die Gruppenunterschiede für jeden Teil des Beurteilungstests einzeln, so sind alle Ergebnisse nicht-signifikant, die Teile B ($p = 0,089$) und D ($p = 0,140$) weisen aber wesentlich geringere Werte auf als die anderen Teile (siehe Tab. 36). Es kann also davon ausgegangen werden, dass es keine signifikanten Unterschiede bezüglich der zentralen Tendenz der Personenwerte zwischen den einzelnen Vorerfahrungsstufen gibt. Dieses Ergebnis des Kruskal-Wallis-Tests wird durch den allgemeineren Mediantest bei unabhängigen Stichproben sowie den Jonckheere-Terpstra-Test für alle Teile des Beurteilungstests bekräftigt (Janssen & Laatz, 2017, S. 668–674).

Tab. 36: Ergebnisse des Kruskal-Wallis-Tests für die einzelnen Teile des Beurteilungstests hinsichtlich der fünf Vorerfahrungsstufen

	BT_A ORM	BT_B ORM	BT_C ORM	BT_D ORM	BT_E ORM
Chi-Quadrat	2,577	8,059	1,382	6,918	2,922
Df	4	4	4	4	4
Asymptotische Signifikanz	,631	,089	,847	,140	,571

a. Kruskal-Wallis-Test b. Gruppenvariable Vorerf_Gruppe

In einem gruppierten Boxplot-Diagramm kann der Personenparameter (in Abb. 85 für Teil A des Beurteilungstests) für die jeweiligen Stufen der Beurteilungsvorerfahrung graphisch verglichen werden. Man sieht, dass bei Teil A der Median für die Stufen 1 und 3 gleich ist und geringer als für die Stufen 2, 4 und 5. Der durchschnittliche Rang beträgt dabei für Stufe 1 26,29, für Stufe 2 36,77, für Stufe 3 28,89, für Stufe 4 29,21 und für Stufe 5 30,12 (siehe Tab. 36). Abgesehen von Stufe 2 nimmt demnach die Rangziffer kontinuierlich zu, was vermuten lässt, dass ein Unterschied in der Beurteilungsleistung zwischen den vier Vorerfahrungsstufen besteht. Diese Unterschiede werden aber gemäß Tab. 36 als nicht-signifikant getestet (für Teil BT_A ORM mit einer asymptotischen Signifikanz von 0,631, das die mit $\alpha = 0,05$ vorgegebene Irrtumswahrscheinlichkeit übersteigt), weshalb die Nullhypothese (H_0 : es gibt für die einzelnen Vorerfahrungsstufen keinen Unterschied in der Beurteilungsleistung) angenommen wird. Wenn die Nullhypothese abgelehnt wird, werden die post-hoc Mehrvergleiche nicht mehr ausgeführt und müssen manuell über den Mann-Whitney-U-Test für die einzelnen Stufen zueinander berechnet werden. Die einzelnen Rangziffern der Beurteilungsteile werden in Tab. 37 zusammengefasst.



Tab. 37: Mittlerer Rang der einzelnen Vorerfahrungsstufen für die Teile des Beurteilungstests

	Stufe 1	Stufe 2	Stufe 3	Stufe 4	Stufe 5
BT_A_ORM	26,29	36,77	28,89	29,21	30,12
BT_B_ORM	31,46	32,91	21,94	36,24	19,25
BT_C_ORM	31,18	30,55	24,00	31,21	31,38
BT_D_ORM	33,64	21,41	27,72	28,76	40,62
BT_E_ORM	30,29	28,23	22,83	34,56	30,31
BT_ABCDE_ORM	33,11	28,59	20,78	33,18	30,12
BT_AB	29,50	35,00	23,61	33,44	23,88
BT_CDE	33,14	25,68	21,89	33,18	32,81

In einem Post-Hoc-Test wird zusätzlich untersucht, welche Gruppen sich im Einzelnen paarweise voneinander unterscheiden. Zu diesem Zweck wurde für jede Gruppenkombination ein entsprechender Mann-Whitney-U-Test mit SPSS durchgeführt, dessen Ergebnis in Tab. 38 zusammengefasst ist. Ein signifikanter Gruppenunterschied ist bei Aufgabenteil B zwischen den Testpersonen der Stufen 3 und 4 ($p = 0,049$) sowie 4 und 5 ($p = 0,033$) festzustellen, doch auch die Vergleiche Stufe 1-Stufe 3 ($p = 0,146$), Stufe 1-Stufe 5 ($p = 0,082$), Stufe 2-Stufe 3 ($p = 0,139$) und Stufe 2-Stufe 5 ($p = 0,085$) liegen relativ nah an der Signifikanzgrenze. Die signifikanten Werte kommen vor allem durch die schwachen Personenmittelwerte der Testpersonen von Stufe 3 und Stufe 5 zustande. Der Dunn-Bonferroni-Tests errechnet für die Gruppenunterschiede zwischen Stufe 3 und Stufe 4 bei Teil B eine nach Cohen (1992, S. 156) mittlere Effektstärke von $r = 0,39$ (für $z = -1,968$ und $N = 26$) und zwischen Stufe 4 und 5 ebenfalls einen mittlerer Effekt von $r = 0,43$ (für $z = -2,138$ und $N = 25$). Für Testteil D kann ein signifikanter Unterschied zwischen der Stufe 2 und der Stufe 5 festgestellt werden ($p = 0,017$), wobei die Effektstärke von $r = 0,55$ (für $N = 19$ und $z = -2,390$) einem starken Effekt entspricht. Der hohe Mittelwert der Gruppe 5 führt dazu, dass auch die Vergleiche mit Stufe 3 ($p = 0,085$) und Stufe 4 ($p = 0,085$) zu geringen Signifikanzwerten führen. Ebenso trifft dies auf den Vergleich Stufe 1-Stufe 2 ($p = 0,107$) sowie für Aufgabe E auf Stufe 3-Stufe 4 ($p = 0,090$) zu.

Tab. 38: Asymptotische Signifikanzwerte (2-seitig) des Mann-Whitney-U-Tests für den Gruppenvergleich bezüglich verschiedener Variablen

	BT A	BT B	BT C	BT D	BT E	BT_ABC		
						DE	BT_AB	BT_CDE
Stufe 1-Stufe 2	,122	,801	,888	,107	,639	,510	,419	,271
Stufe 1-Stufe 3	,847	,146	,300	,425	,325	,114	,352	,130
Stufe 1-Stufe 4	,625	,361	,984	,432	,435	,984	,433	,984
Stufe 1-Stufe 5	,500	,082	,915	,489	,973	,632	,386	,973
Stufe 2-Stufe 3	,385	,139	,386	,350	,420	,286	,133	,516
Stufe 2-Stufe 4	,242	,564	,942	,229	,353	,464	,868	,211
Stufe 2-Stufe 5	,300	,085	,932	,017	,868	,869	,139	,428
Stufe 3-Stufe 4	,978	,049	,306	,890	,090	,092	,232	,128
Stufe 3-Stufe 5	,805	,656	,404	,085	,468	,207	,882	,227
Stufe 4-Stufe 5	,952	,033	1,000	,085	,658	,681	,228	,953

Der Zusammenhang zwischen den Resultaten des Beurteilungstests und der Stufe der Vorerfahrungen, der die Testpersonen zugeordnet wurden, zeigt, dass man die Ergebnisse differenziert beurteilen muss. Betrachtet man die asymptotischen Signifikanzwerte für die Unterschiede der einzelnen Vorerfahrungsstufen zueinander (siehe Tab. 38), so sind bezüglich des Personenwertes des Gesamtbeurteilungstests (BT_ABCDE) alle Werte nicht-signifikant, es fällt jedoch auf, dass die eindeutig geringsten Werte der verschiedenen Paare immer in Kombination mit der Vorerfahrungsstufe 3 auftreten (Stufe 1-Stufe 3 mit $p = 0,114$, Stufe 2-Stufe 3 mit $p = 0,286$, Stufe 3-Stufe 4 mit $p = 0,092$ und Stufe 3-Stufe 5 mit $p = 0,207$). Die p -Werte sind zwar nicht signifikant, aber es zeigt sich, dass die Probandinnen und Probanden der Stufe 3 (Vorerfahrungen in Sportarten, in denen Schlag(wurf)bewegungen vorkommen) im Vergleich zu Personen mit keinerlei Beurteilungserfahrung (Stufe 1) sowie zur Gruppe der sehr Beurteilungserfahrenen (Stufe 5) auffällig niedrigere Signifikanzwerte haben als die anderen Kombinationen. Bezieht man die Mittelwerte des gesamten Beurteilungstests (Variable BT_ABCDE) in die Interpretation mit ein (siehe Tab. 35: Stufe 3 mit $\theta = 0,864$ im Vergleich zu den Stufen 1 mit $\theta = 1,135$ oder 4 mit $\theta = 1,131$), so lässt sich die Annahme formulieren, dass Testpersonen der Gruppe 3 beim Beurteilungstest tendenziell schlechter abschneiden als diejenigen der anderen Gruppen. Während die Unterschiede zwischen Gruppe 3 und Gruppe 5 durch die größere Beurteilungserfahrung auf Stufe 5 erklärbar sind, trifft diese Tatsache auf das bessere Abschneiden von Gruppe 1 im Vergleich zu Gruppe 3 nicht zu.

Betrachtet man die Ausprägung des Personenwerts für die einzelnen Beurteilungsaufgaben, so gibt es für den Beurteilungserfolg von Teil A zwar keinen Anlass zur Annahme von signifikanten Gruppenunterschieden, der hohe erreichte Personenmittelwert für Teil A von Vorerfahrungsstufe 2 ($\theta = 2,057$) ist jedoch überraschend und führt zu niedrigen Signifikanzwerten bei Vergleichen mit Stufe 1 ($p = 0,122$) oder Stufe 4 ($p = 0,242$). Bei dieser einfachsten der Beurteilungsaufgaben, dem Ordnen der Bewegungen nach der qualitativen Ausführung einzelner Bewegungsmerkmale, spielt es eine untergeordnete Rolle, ob bereits schlagwurfähnliche Bewegungen beurteilt wurden, allein die Tatsache, dass Beurteilungserfahrung vorliegt, bedingt ein besseres Testergebnis als bei Personen, die sich als völlig beurteilungsun erfahren einstufen ($\theta = 1,448$).

Bei Teil B gibt es mehrere niedrige Signifikanzwerte, die interpretiert werden müssen. Testpersonen der Vorerfahrungsstufe 4 waren mit einem Personenmittelwert von $\theta = 1,559$ ($SD = 0,638$) in diesem Aufgabenteil signifikant besser als ihre Kommilitoninnen und Kommilitonen auf Stufe 3 ($p = 0,015$) und Stufe 5 ($p = 0,015$). Durch die sehr niedrige Personeneigenschaft der Testpersonen auf Stufe 3 ($\theta = 1,137$, $SD = 0,467$) ergeben sich auch auffällige, knapp das 0,05-Niveau überschreitende Gruppenunterschiede mit Stufe 1 ($p = 0,078$) und Stufe 2 ($p = 0,080$). Auch das schlechte Abschneiden der Testpersonen auf Stufe 5 ($\theta = 1,084$, $SD = 0,470$) führt zu knapp nicht-signifikanten Unterschieden mit Stufe 1 ($p = 0,058$) und Stufe 2 ($p = 0,078$). Als Erklärung für die niedrige Personenfähigkeit der Gruppenmitglieder von Stufe 3 kann wieder die vorgeprägte, aber verunsichernde Beurteilungserfahrung von anderen Schlag(wurf)bewegungen dienen, während die Testpersonen von Stufe 5 möglicherweise aufgrund ihrer bereits vorhandenen Beurteilungsmuster Einschätzungen vornehmen, die zwar ihrer Beurteilungsvorerfahrung entsprechen, nicht aber unbedingt kohärent mit den vorgegebenen Qualitätsstufen sind. Dies kann beispielsweise der Fall sein, wenn eine Testperson mit gut ausgebildeten Athleten arbeitet, im Beurteilungstest aber relativ defizitäre Bewegungsausführungen von Schülern einschätzen muss. Durch die Konzentration auf ein einzelnes und bereits zur Beurteilung vorgegebenes Bewegungsmerkmal verstärkt sich möglicherweise der Eindruck, dass dieses schlechter ausgeführt wurde als es bei einer ungesteuerten Bewegungsbeobachtung wahrgenommen werden würde.

Die Personenmittelwerte von Teil C unterscheiden sich nicht so stark, dass eine begründete Annahme über den Einflussfaktor der Beurteilungserfahrung vorgenommen werden könnte. Der schlechte Personenmittelwert von Stufe 3 ($\theta = 0,609$, $SD = 0,742$) kann möglicherweise wieder auf die Beurteilungsunsicherheit durch die Beurteilungserfahrung schlagwurfähnlicher Bewegungen aus anderen Sportarten zurückgeführt werden. Dafür zeigt Teil D mit $p = 0,017$ einen signifikanten Gruppenunterschied zwischen der am Besten in diesem Testteil abschneidenden Gruppe 5 ($\theta = 1,605$, $SD = 0,762$) und den am schlechtesten abschneidenden Gruppenmitgliedern der Stufe 2 ($\theta = 0,801$, $SD = 0,546$). Auch bei den Stufen 3 ($p = 0,085$) und 4 ($p = 0,085$) deutet sich ein Unterschied zu Stufe 5 an, der knapp an der 0,05-Signifikanzgrenze liegt, was wiederum an den deutlich schlechteren Personenmittelwerten von Stufe 3 ($\theta = 0,998$, $SD = 0,556$) und Stufe 4 ($\theta = 1,076$, $SD = 0,610$) liegt. Abgesehen vom wiederum überraschend guten Abschneiden der Testpersonen von Stufe 1 ($\theta = 1,405$, $SD = 0,915$) kann dieser Aufgabenteil umso besser gelöst werden, je größer die Erfahrung beim Beurteilen sportlicher Bewegungen im Allgemeinen und der leichtathletischen Wurfbewegung im Speziellen ist. Eine mögliche Erklärung liegt darin, dass diese Beurteilungsaufgabe am ehesten der beim Bewegungslernen zur Anwendung kommenden morphologischen Bewegungskorrektur entspricht, indem defizitär ausgeführte Bewegungsmerkmale beobachtet und für Rückmelde- oder Beurteilungsprozesse identifiziert werden. In dieser Beobachtungssituation sind Testpersonen mit höherer Beurteilungserfahrung geübter als unerfahrene Beurteiler, weil sie ein bekanntes Beurteilungsmuster anwenden und somit die komplexe Beurteilungsaufgabe der Fehleranalyse zuverlässiger lösen können.

Für Beurteilungsteil E lassen sich anhand der asymptotischen Signifikanzwerte keine eindeutigen Tendenzen erkennen, ob die Beurteilungsvorerfahrung hilfreich ist für die Bewältigung der Aufgabe. Lediglich der Vergleich zwischen Stufe 3 und Stufe 4 zeigt einen niedrigen Wert von $p = 0,090$. Das qualitative Einschätzen aller sechs Bewegungsmerkmale in einem Beobachtungsdurchgang in Echtzeit stellt möglicherweise auch für erfahrene Beurteiler eine Überforderung dar, so dass deren Erfahrungsvorsprung keine nachweisbare Ursache für Gruppenunterschiede darstellt. Diese Vermutung wird auch durch die Tatsache gestützt, dass die Gruppenmitglieder von Stufe 1 ($\theta = 1,139$, $SD = 0,936$), welche die geringste Vorerfahrung in der Bewegungsbeurteilung haben, zusammen mit Stufe 4 ($\theta = 1,144$, $SD = 0,565$) den höchsten Personenmittelwert bei dieser Aufgabenstellung erreichen.

Es lässt sich zusammenfassen, dass vor allem die guten Personenmittelwerte der Stufe 1 (Testpersonen ohne Beurteilungsvorerfahrung) abgesehen von Teil A bei allen anderen Beurteilungsaufgaben auffallen. Ebenso überrascht, dass die Testpersonen der Stufe 3 (mit Beurteilungsvorerfahrung ausschließlich in Sportarten mit schlag(wurf)ähnlichen Bewegungen wie Volleyball oder Handball) bei allen Beurteilungsteilaufgaben im Vergleich zu den anderen Stufen unterdurchschnittliche Personenmittelwerte erzielen. Der Erfahrungsvorsprung der Stufe 5 gegenüber Stufe 4, der sich in der vorhandenen Beurteilungserfahrung leichtathletischer Schlagwurfbewegungen begründet, führt nur bei Aufgabe D zu auffällig besseren Personenmittelwerten, während bei Aufgabenteil B die leichtathletische Vorerfahrung sogar leistungsmindernd wirkt und die Gruppenmitglieder der Stufe 4 signifikant besser abschneiden. Bei Testteil D kann die hohe und spezifische Beurteilungsvorerfahrung von Stufe 5 als leistungsbestimmender Faktor gegenüber den Stufen 2, 3 und 4 angenommen werden.

5.1.5 Eigenrealisation

Der ermittelte durchschnittliche Personenparameter der Eigenrealisation der Schlagwurfbewegung beträgt für die Probandengruppe $\theta = 0,390$ (SD = 0,380). Die auftretenden Häufigkeiten der einzelnen Werte (siehe Abb. 86) stellen nach rechnerischer Überprüfung mit dem Shapiro-Wilk-Test ($p = 0,792$) keine Normalverteilung dar. Eine genaue Erklärung, wie diese Personenfähigkeitswerte zustande kommen und wie sich die einzelnen Beurteilungskriterien statistisch bewährt haben, findet sich in Kapitel 5.2.3.

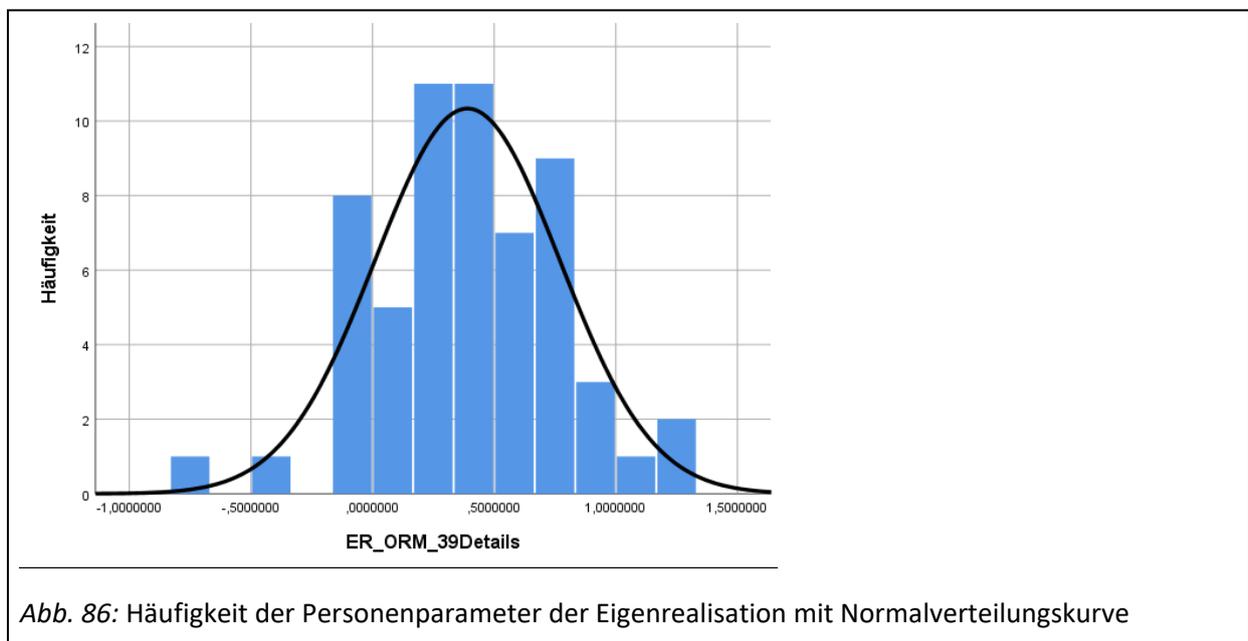


Abb. 86: Häufigkeit der Personenparameter der Eigenrealisation mit Normalverteilungskurve

Abschließend stellt sich wieder die Frage, ob ein nachweisbarer Zusammenhang zwischen der Qualität der Eigenrealisation und der gemessenen Beurteilungsleistung besteht. Wenn man die Personenwerte der 39-Items-Detailanalyse als Variable für die Eigenrealisation nimmt, so kommt man zu überraschend deutlichen Ergebnissen. Während der Korrelationskoeffizient mit Teil A ($r = 0,051$ mit $p = 0,530$) keinen Zusammenhang mit der Eigenrealisation anzeigt, korreliert Teil B mit $r = -0,487$ (mit $p = 0,000$) auf dem

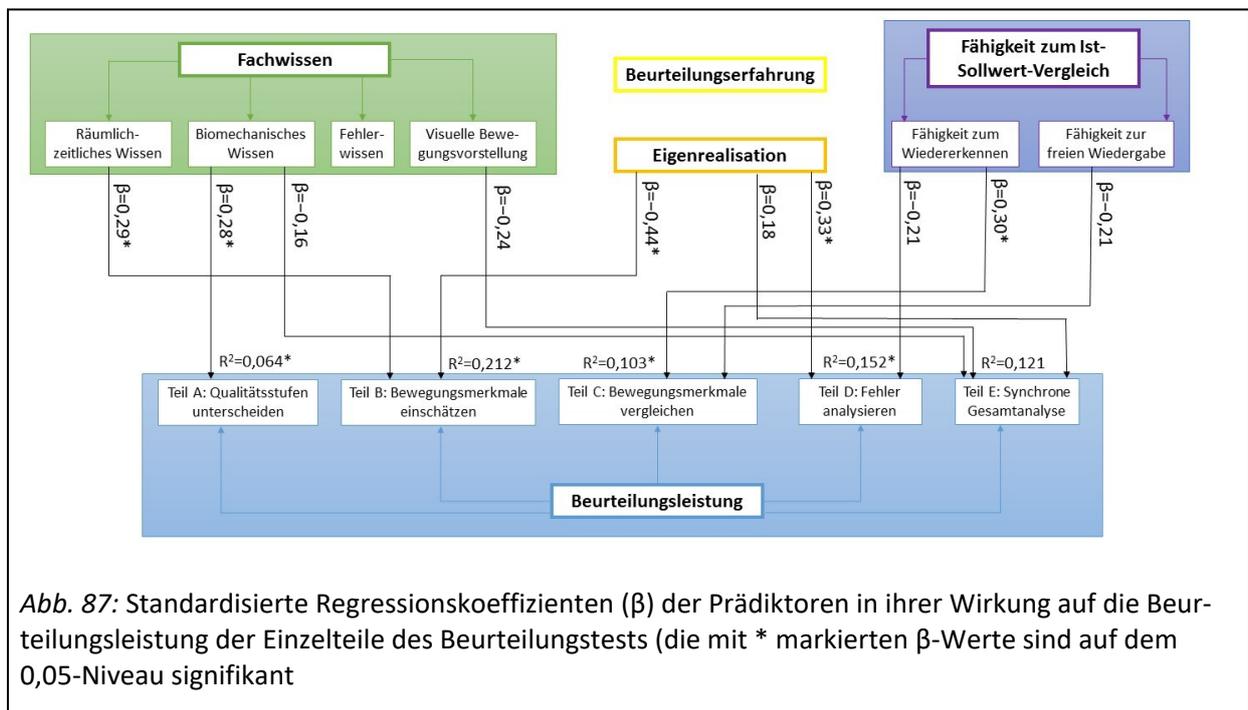
0,01-Niveau signifikant negativ mit der Demonstrationsfähigkeit der Schlagwurfbewegung. Da Teil A eine sehr einfache Zuordnungsaufgabe darstellt, die von vielen Testpersonen sehr gut gelöst wurde, ist die fehlende Bedeutung der Eigenrealisation durchaus nachvollziehbar. Dass Testpersonen mit einer guten Eigenrealisation jedoch bei Beurteilungsaufgabe B deutlich schlechter abschneiden als solche mit einem niedrigen Personenwert, ist überraschend, zumal das richtige Identifizieren der Qualitätsstufe eines einzelnen Bewegungsmerkmals noch keine allzu komplexe Aufgabe darstellt.

Die komplexeren Beurteilungsaufgaben korrelieren durchwegs positiv mit der Eigenrealisation. Vor allem für die Fehleranalyse von Teil D ist mit einem Korrelationskoeffizienten von $r = 0,376$ ($p = 0,003$) der Zusammenhang mit der Eigenrealisation auf einem 0,01-Signifikanzniveau besonders deutlich, aber auch die Teile C ($r = 0,185$ mit $p = 0,160$) und E ($r = 0,179$ mit $p = 0,176$) weisen einen schwach positiven Zusammenhang auf. Diese Zusammenhänge deutet an, dass ein gutes Bewegungskönnen vorteilhaft für die komplexen Fremdbeurteilungsaufgaben wie die eines Bewegungsvergleichs, einer Fehleranalyse oder einer synchronen Gesamtanalyse derselben Bewegung sein kann.

5.1.6 Zusammenfassung: Mögliche Wirkungen der Einflussfaktoren auf die Beurteilungsleistung

Da Korrelationskoeffizienten lediglich eine Aussage über die Stärke des Zusammenhangs, nicht jedoch über seine Richtung zulassen, wird in einem letzten Schritt eine multiple Regressionsanalyse für die jeweiligen Teile des Beurteilungstests durchgeführt. Es wird daher auf Grundlage der berechneten Korrelationsbeziehungen überprüft, ob und wie stark die untersuchten unabhängigen Variablen Fachwissen, allgemeine Bewegungsvorstellung, Beurteilungserfahrung und Eigenrealisation die abhängige Variable (in diesem Fall die entsprechenden Personenparameter der Einzelteile des Beurteilungstests) beeinflussen. Hierbei erweisen sich vor allem die Beurteilungsaufgaben der Fehleranalyse (Teil D) und der Qualifizierung einzelner Bewegungsmerkmale (Teil B) als gelungen, weil aus der Korrelationsmatrix mehrere Prädiktoren identifiziert werden können, welche die Vorhersage der Beurteilungsleistung beeinflussen könnten. In der Regressionsanalyse kann festgestellt werden, dass speziell die Qualität der Eigenrealisation und der TKBV-Teil1 (Fähigkeit zum Wiedererkennen) mit $R^2 = 0,152$ einen signifikant nachweisbaren positiven Einfluss auf die Beurteilungsleistung bei Aufgaben der Fehleranalyse (Teil D) ausübt (siehe Kapitel 5.1.6.1). Die Beurteilungsfähigkeit bei der Qualifizierung einzelner Beurteilungsmerkmale (Teil B) muss differenzierter betrachtet werden: es kann eine negative Wirkung der Eigenrealisation und eine positive Wirkung des räumlich-zeitlichen Fachwissens (mit $R^2 = 0,212$) signifikant nachgewiesen werden (siehe Kapitel 5.1.6.2). Für die Fähigkeit zur Unterscheidung der einzelnen Qualitätsstufen (Teil A) konnte lediglich das biomechanische Fachwissen (mit $R^2 = 0,076$) als signifikanter Einflussfaktor identifiziert werden (siehe Kapitel 5.1.6.3). Die Testleistung beim Vergleichen von Bewegungsmerkmalen bei zwei verschiedenen Bewegungsausführungen (Teil C) kann mithilfe der zwei Teile des TKBV erklärt werden (siehe Kapitel 5.1.6.4), während für die synchrone Gesamtanalyse das biomechanische Fachwissen, die visuelle Bewegungsvorstellung und die Eigenrealisation als Prädiktoren ermittelt werden konnten, wenngleich ohne das 0,05-Signifikanzniveau zu erreichen (siehe Kapitel 5.1.6.5).

Zusammenfassend wird in Abb. 87 das in Kapitel 3 entwickelte Grundmodell der Beurteilungsleistung und ihrer möglichen Einflussfaktoren (siehe Abb. 30) entsprechend den Erkenntnissen dieser Studie modifiziert und die ermittelten Kausalitätsbeziehungen mittels der standardisierten Regressionskoeffizienten (β) dargestellt. Der R-Quadrat-Wert bei den einzelnen Beurteilungsteilen zeigt, dass lediglich leichte bis mittlere Effektstärken nachgewiesen werden konnten, die aber abgesehen von Teil E immerhin das 0,05-Signifikanzniveau erfüllen. Somit müssen die Schlussfolgerungen, die aus den Ergebnissen und Einzelwerten gezogen werden, vorsichtig und zurückhaltend formuliert werden, wenn gleich gewisse Zusammenhänge erkennbar werden, die in einer Folgestudie nochmals genauer untersucht werden könnten. Vor allem die geringe Zahl der Testpersonen, die den Ergebnissen zugrunde liegt, sowie die relativ homogene Zusammensetzung der Testgruppe, verringern die Wahrscheinlichkeit, nachgewiesene Effekte auch in der Grundgesamtheit zu entdecken.



Bei Regressionsberechnungen muss auch der Tatsache Rechnung getragen werden, dass die Prädiktoren untereinander korrelieren, denn je stärker diese Interkorrelation ausfällt, desto weniger kann ein Prädiktor unabhängig von anderen Prädiktoren die Unterschiede im Kriterium erklären. Betrachtet man die Korrelationsbeziehungen der möglichen Einflussfaktoren untereinander, so sind die Interkorrelationen insgesamt unauffällig. Die einzig signifikante Korrelation besteht zwischen dem räumlich-zeitlichen und dem biomechanischen Fachwissen ($r = 0,328$, $p = 0,011$), wobei beide Variablen gleichzeitig in keiner Beurteilungsaufgabe als Prädiktoren nachgewiesen werden konnten. Die visuelle Bewegungsvorstellung (Bildsortieraufgabe), welche eine leicht negative Korrelation mit der Eigenrealisation ($r = -0,149$, $p = 0,260$) und eine leicht positive mit dem TKBV-Teil1 ($r = 0,161$, $p = 0,224$) zeigt, erweist sich lediglich als nicht signifikanter Prädiktor der synchronen Gesamtanalyse (Teil E). Das räumlich-zeitliche Fachwissen stellt sich als problematischste Variable heraus, weil es beispielsweise bei der Regressionsanalyse von Teil D des Beurteilungstests mit allen in Frage kommenden Prädiktoren Fehlerwissen ($r = 0,220$, $p = 0,094$), Bildreihe ($r = 0,154$, $p = 0,243$), Eigenrealisation ($r = 0,234$, $p = 0,075$) und TKBV-Teil 1 ($r = 0,209$, $p = 0,112$) positiv korreliert. Ähnliches trifft für das biomechanische Fachwissen zu, welches mit der Bildsortierleistung ($r = 0,158$, $p = 0,233$) und TKBV-Teil 1 ($r = 0,250$,

$p = 0,056$) zusammenhängt (siehe Tab. 54 im Appendix für die möglichen Prädiktoren von Teil D des Beurteilungstests). Daher wird bei den Regressionsanalysen besonders darauf geachtet, die spezifischen Varianzanteile der Prädiktoren, also der Anteil, den jeder Prädiktor ohne die anderen Prädiktoren mit dem Kriterium gemeinsam hat, zu berücksichtigen, indem das standardisierte Regressionsgewicht β interpretiert wird. Mit diesem Wert kann nämlich ausgedrückt werden, dass sich der Wert des Kriteriums durchschnittlich um β -Standardabweichungen ändert, wenn sich der Wert des Prädiktors um eine Standardabweichung erhöht und gleichzeitig alle anderen Prädiktoren konstant gehalten werden.

5.1.6.1 Einflussfaktoren auf die Fehleranalyse

Betrachtet man demnach die Korrelationsbeziehung der Beurteilungsleistung von Teil D mit den möglichen Einflussvariablen, so fällt mit $r = 0,376$ ($p = 0,003$) die signifikant positive Beziehung zur Eigenrealisation auf. Auch das Fehlerwissen zeigt mit $r = 0,183$ ($p = 0,166$) eine ähnliche Tendenz, während die Bildsortierleistung mit $r = -0,174$ ($p = 0,187$) und Teil 2 des TKBV mit $r = -0,192$ ($p = 0,146$) negativ mit der Fähigkeit zur Fehleranalyse korrelieren. Führt man nun zunächst eine multiple Regressionsanalyse mit den vier ausgewählten Prädiktoren Fehlerwissen, Bildsortierleistung, Eigenrealisation und TKBV-Teil1 mit der Einschluss-Methode durch, so wird ein multipler Korrelationskoeffizient von $R = 0,432$ für den gemeinsamen Zusammenhang mit dem Kriterium errechnet, wobei Unterschiede in den vier Prädiktoren 18,7 Prozent der Unterschiede in der Beurteilungsleistung vorhersagen ($R^2 = 0,187$). Die Güte der Vorhersage lässt sich über den Standardschätzfehler ermitteln und beträgt 0,678 (d. h. dass der Fehler bei der Schätzung der Beurteilungsleistung im Mittel mehr als 0,678 Logit beträgt). In einer anschließenden Varianzanalyse zeigt sich, dass der durch die vier Prädiktoren vorhergesagte Anteil der Kriteriumsvarianz (R^2) mit $p = 0,023$ ($F = 3,106$) signifikant von Null verschieden ist. Betrachtet man den Regressionskoeffizienten mit $B = 0,611$, so kann man daraus schlussfolgern, dass sich die Beurteilungsleistung um 0,611 Logit erhöht, wenn sich die Qualität der Eigenrealisation um ein Logit verbessert. Vergleicht man die standardisierten Regressionsgewichte der Prädiktoren, so ergibt sich für die Eigenrealisation der größte Wert von $\beta = 0,320$ ($p = 0,012$). Dies besagt, dass sich der vorhergesagte Wert der Beurteilungsleistung signifikant um durchschnittlich 0,320 Standardabweichungen ändert, wenn sich der Wert der Eigenrealisation um eine Standardabweichung verändert und gleichzeitig die drei anderen Prädiktoren kontrolliert werden. Alle anderen Prädiktoren besitzen einen kleineren Beta-Koeffizienten und sind nicht statistisch signifikant. Betrachtet man die Semipartialkorrelation ($r_{sp} = 0,318$) für die Eigenrealisation, welche nur geringfügig kleiner ist als die Produkt-Moment-Korrelation ($r = 0,343$), so kann man schlussfolgern, dass zwischen der Eigenrealisation und der Beurteilungsleistung weiterhin ein Zusammenhang besteht, auch wenn für den Prädiktor der Eigenrealisation die anderen Prädiktoren kontrolliert werden. Ein Konditionsindex von 20,901 für Dimension 5 spricht dafür, von einer mäßigen Kollinearität (Konditionsindex zwischen 15 und 30) zwischen den Prädiktoren auszugehen. Das Regressionsmodell weist eine mittlere Effektstärke auf ($f^2 = 0,230$).

Auch mit der hierarchisch linearen multiplen Regressionsanalyse wird für Teil D die Eigenrealisation mit $p = 0,011$ als signifikante Prädiktorvariable ermittelt. Der Personenwert des TKBV-Teil 1 erweist sich mit $p = 0,097$ ebenfalls als ernst zu nehmende Einflussvariable auf die Beurteilungsleistung von Teil D, während sich die anderen Prädiktoren im Rahmen der Rückwärtselimination als eindeutig nicht-signifikant erweisen. Dieses 2-Variablen-Modell erweist sich im F-Test ($F(1,57) = 5,367$, $p = 0,007$) als Ganzes signifikant. Der t-Test für den Regressionskoeffizienten der Eigenrealisation ist signifikant

($t = 2,689$, $p = 0,009$), derjenige des TKBV-Teil 1 hingegen liegt leicht über dem 0,05-Signifikanzniveau ($t = -1,705$, $p = 0,094$). Durch das negative Vorzeichen des t-Koeffizienten bei TKBV-Teil 1 deutet sich an, dass mit einem zunehmenden Personenwert des TKBV-Teil1 um 1 Logit die Beurteilungsleistung in Teil D um 0,150 Logit sinkt. Steigt hingegen der Personenwert der Eigenrealisation um 1 Logit, so nimmt auch der Personenwert des Beurteilungsteils D um 0,656 Logit zu. Das korrigierte R-Quadrat beträgt 0,152, wodurch 15,2 Prozent der Gesamtstreuung in der Beurteilungsleistung von Teil D durch die Variablen Eigenrealisation ($\beta = 0,329$) und TKBV-Teil1 ($\beta = -0,208$) erklärt werden kann. Die Effektstärke beträgt $f^2 = 0,179$, was einem mittleren Effekt entspricht.

5.1.6.2 Einflussfaktoren auf die Beurteilung einzelner Bewegungsmerkmale

Bei den Korrelationsbeziehungen der untersuchten Einflussfaktoren mit Teil B des Beurteilungstests nach dem ordinalen Rasch-Modell fällt auf, dass die Eigenrealisation mit $r = -0,487$ ($p = 0,000$) signifikant negativ mit der Beurteilungsleistung korreliert. Abgesehen von der Vorerfahrung ($r = -0,097$, $p = 0,463$) weisen alle anderen Variablen sehr geringe und somit vernachlässigbare Korrelationskoeffizienten auf. Etwas anders sieht das Ergebnis aus, wenn man Teil B nach dem dichotomen Rasch-Modell auswertet. Die Eigenrealisation bleibt mit $r = -0,430$ ($p = 0,001$) der entscheidende Faktor, aber auch das räumlich-zeitliche Fachwissen mit $r = 0,129$ ($p = 0,332$), das Fehlerwissen mit $r = 0,171$ ($p = 0,194$), die Bildreihe mit $r = 0,173$ ($p = 0,189$) und die Beurteilungsvorerfahrung mit $r = -0,118$ ($p = 0,375$) weisen beachtenswerte Korrelationskoeffizienten auf (siehe Tab. 55). Daher wird für die Regressionsanalyse als Kriterium der Personenwert der Beurteilungsleistung nach dem dichotomen Rasch-Modell bevorzugt, weil hier die Prädiktorenausprägung den theoretischen Vorüberlegungen besser entspricht als nach dem ordinalen Rasch-Modell. Eine Regressionsanalyse wird aber für beide Auswertungsverfahren durchgeführt, um die Ergebnisse miteinander vergleichen zu können.

Wenn man die Regression für alle fünf Prädiktoren gemeinsam (nach dem Einschluss-Verfahren) berechnet, so lassen sich bei einer mittleren Effektstärke ($f^2 = 0,303$) 23,3 Prozent der Beurteilungsleistung mit Veränderungen in der Eigenrealisation, dem räumlich-zeitlichen Fachwissen, dem Fehlerwissen, der Bildreihe und der Beurteilungsvorerfahrung erklären. Dieses Modell ist insgesamt signifikant ($F = 3,233$, $p = 0,013$), wobei lediglich die Eigenrealisation mit $p = 0,001$ deutlich unterhalb und das räumlich-zeitliche Fachwissen mit $p = 0,062$ knapp oberhalb der 0,05-Signifikanzgrenze liegen. Mit einem Konditionsindex von 18,409 bei Dimension 6 ist von einer mittleren Kollinearität zwischen den Prädiktoren auszugehen, weshalb über die Eliminierung von kollinearen Prädiktoren nachgedacht werden muss. Mit der Rückwärtsmethode lässt sich ermitteln, dass das plausibelste Regressionsmodell lediglich die Eigenrealisation mit einem standardisierten Beta-Koeffizienten von $\beta = -0,444$ ($p = 0,001$) und das räumlich-zeitliche Fachwissen mit $\beta = 0,287$ ($p = 0,023$) als signifikante Prädiktoren umfasst. Das R-Quadrat für das Modell mit der Eigenrealisation allein beträgt dann $R^2 = 0,135$, für das Modell mit Eigenrealisation und räumlich-zeitlichem Fachwissen erhöht sich der Wert auf $R^2 = 0,212$. Somit können 21,2 Prozent der Beurteilungsleistung von Teil B mit diesen beiden Prädiktoren, deren Interkollinearität gering ist (Konditionsindex von 0,298), vorhergesagt werden. Die Beurteilungsleistung steigt dabei um 0,278 Logit, wenn sich das räumlich-zeitliche Fachwissen um ein Logit erhöht, und sinkt um 0,924 Logit, wenn sich die Eigenrealisation um ein Logit verbessert.

Bei einer hierarchisch angelegten Regressionsanalyse mit der ordinal ausgewerteten Beurteilungsleistung als Kriterium, bei der zuerst die Eigenrealisation allein (Modell 1) und dann die Eigenrealisation und die Vorerfahrung (Modell 2) als Prädiktoren festgelegt werden, erweist sich Modell 1 mit $R^2 = 0,177$

(Standardschätzfehler 0,532) als durchaus aussagekräftig, denn 17,7 Prozent der Beurteilungsleistung können über die Qualität der Eigenrealisation vorhergesagt werden. Hier wird der Eigenrealisation eine stärkere Wirkung bestätigt als beim dichotomen Auswertungsverfahren. Die Hinzunahme der Beurteilungsvorerfahrung als Prädiktor ($\beta = -0,025$) kann jedoch keine Verbesserung mehr erreichen ($R^2 = 17,8$ für Modell 2). Der F-Test bestätigt, dass Modell 1 als Ganzes signifikant ist ($F(1,57) = 12,298$, $p = 0,001$), weshalb die Analyse weiter fortgesetzt werden kann. Die t-Tests sowohl für die Regressionsvariable ($t = -3,507$, $p = 0,001$) als auch für die Konstante ($t = 16,568$, $p = 0,000$) fallen signifikant aus. Wenn demnach der Personenwert der unabhängigen Variable Eigenrealisation um 1 Logit steigt, so nimmt der Personenwert des Beurteilungstests von Teil B um 0,645 Logit ab. Das Regressionsmodell mit der Eigenrealisation als einzigem Prädiktor erweist sich dabei mit einem standardisierten Beta-Koeffizienten von $\beta = -0,421$ als signifikant, eine Kollinearität zwischen den Variablen gilt mit einem Konditionsindex von 2,476 als sehr unwahrscheinlich. Mit einer Effektstärke nach Cohen von $f^2 = 0,215$ kann man von einem mittleren Effekt sprechen.

5.1.6.3 Einflussfaktoren auf die Unterscheidung der Qualitätsstufen

Die Korrelationsbeziehungen der Beurteilungsleistung bei der Unterscheidung der Qualitätsstufen (Teil A) mit den möglichen Einflussfaktoren zeigt, dass lediglich das biomechanische Fachwissen mit $r = 0,290$ ($p = 0,026$), das Fehlerwissen mit $r = 0,129$ ($p = 0,332$) und die Ergebnisse des TKBV-2 mit $r = 0,142$ ($p = 0,284$) einen beachtenswerten Zusammenhang zeigen (siehe Tab. 56). Auch hier wird wieder das dichotome dem ordinalen Auswertungsverfahren vorgezogen, wenngleich die Unterschiede in den Korrelationskoeffizienten sehr gering ausfallen.

Bei der Regressionsanalyse nach dem Einschussverfahren lässt sich für die drei Prädiktoren ein R-Quadrat von $R^2 = 0,093$ errechnen. Dieses Modell ist jedoch mit $p = 0,142$ ($F = 1,889$) nicht signifikant. Wenn man eine Rückwärtselimination durchführt, so ergibt sich lediglich bei dem Modell mit dem biomechanischen Fachwissen ein signifikanter Wert von $p = 0,035$ mit einem R-Quadrat von $R^2 = 0,076$. Nimmt man das Fehlerwissen noch mit dazu, würde R-Quadrat auf $R^2 = 0,089$ steigen, jedoch wird das 2-Prädiktoren-Modell mit $p = 0,073$ dann nicht-signifikant. Der standardisierte Beta-Koeffizient des biomechanischen Fachwissens beträgt dann $\beta = 0,275$ ($p = 0,035$). Wenn der Personenparameter des biomechanischen Fachwissens um ein Logit steigt, dann nimmt der Personenparameter von Teil A des Beurteilungstests um 0,539 Logit zu. Es darf also durchaus angenommen werden, dass das biomechanische Fachwissen einen gewissen Einfluss auf die Beurteilungsleistung von Teil A hat, den Bewegungsbeispielen die richtigen Qualitätsstufen zuzuordnen. Somit lassen sich 7,6 Prozent der Beurteilungsleistung mit dem biomechanischen Fachwissen erklären. Mit einer berechneten Effektstärke von $f^2 = 0,082$ entspricht dies einem geringen Effekt.

5.1.6.4 Einflussfaktoren auf die Beurteilung des Bewegungsvergleichs

Die Korrelationsmatrix, welche die Zusammenhänge zwischen der Beurteilungsleistung des Bewegungsvergleichs (Teil C des Beurteilungstests) und den möglichen Einflussfaktoren darstellt (siehe Tab. 57), zeigt fünf auffällige Prädiktoren. Die Eigenrealisation mit $r = 0,185$ ($p = 0,160$) und der TKBV-Teil 2 mit $r = 0,231$ ($p = 0,079$) weisen eine positive Beziehung, das räumlich-zeitliche Fachwissen mit $r = -0,157$ ($p = 0,235$), die Bildreihe mit $r = -0,206$ ($p = 0,118$) und der TKBV-Teil1 mit $r = -0,196$ ($p = 0,137$)

hingegen eine negative Beziehung zur Beurteilungsleistung auf, wenngleich alle Korrelationskoeffizienten nicht-signifikant sind.

Führt man eine Regressionsanalyse nach der Standardmethode für die fünf genannten Prädiktoren durch, so erweist sich das gesamte Regressionsmodell als nicht-signifikant ($F = 1,534$, $p = 0,195$), wobei damit nur 12,6 Prozent der Beurteilungsleistung erklärt werden könnten ($R^2 = 0,126$). Überraschend ist hier aber, dass für die beiden Personenwerte des TKBV mit $\beta_{TKBV1} = -0,247$ ($p = 0,080$) und $\beta_{TKBV2} = 0,192$ ($p = 0,164$) die standardisierten Beta-Koeffizienten mit den geringsten Signifikanzwerten auftreten und in entgegengesetzte Richtung deuten. Wendet man die Methode der Rückwärtselimination an, so umfasst das plausibelste Regressionsmodell ebenfalls die beiden Variablen TKBV-Teil1 und TKBV-Teil 2 und erfüllt mit $p = 0,047$ ($F = 3,227$) sogar die Signifikanzgrenze. Mit diesen beiden Prädiktoren ($\beta_{TKBV1} = -0,295$ ($p = 0,027$) und $\beta_{TKBV2} = 0,214$ ($p = 0,106$)) sind immerhin 10,3 Prozent der Beurteilungsleistung des Bewegungsvergleichs vorhersagbar, wobei mit dem Ansteigen des Personenparameterwertes von TKBV-Teil1 um ein Logit die Beurteilungsleistung von Teil C um 0,195 Logit sinkt, während sie bei einem analogen Anstieg von TKBV-Teil2 um 0,118 Logit steigt. Die letztlich aussagekräftigste Variable dieser Beurteilungsaufgabe ist folglich der Personenwert von TKBV-Teil 1. Die Effektstärke für das 2-Prädiktoren-Modell von $f^2 = 0,114$ entspricht nur einem geringen Effekt. Ein Konditionsindex von 4,525 spricht für eine geringe Kollinearität zwischen den beiden Prädiktoren.

5.1.6.5 Einflussfaktoren auf die Beurteilung der synchronen Gesamtanalyse

Beim Vergleich der beiden Auswertungsverfahren erweisen sich die Korrelationsbeziehungen mit der Beurteilungsleistung nach dem dichotomen Rasch-Modell etwas aussagekräftiger als nach dem ordinalen Rasch-Modell (siehe Tab. 58). Das biomechanische Fachwissen mit $r = -0,219$ ($p = 0,096$) und die Bildsortierleistung mit $r = -0,207$ ($p = 0,116$) lassen einen negativen Zusammenhang vermuten, die Eigenrealisation hingegen zeigt mit $r = 0,186$ ($p = 0,159$) eine positive Korrelation.

Das Regressionsmodell mit den drei genannten Variablen verfehlt mit $p = 0,068$ ($F = 2,513$) die Signifikanzgrenze nur knapp, würde aber mit $R^2 = 0,121$ immerhin 12,1 Prozent der Beurteilungsleistung von Teil E erklären, was fast einer mittleren Effektstärke entspricht ($f^2 = 0,138$). Ein Konditionsindex von 9,241 bei nur drei Variablen weist auf eine relativ wahrscheinliche Kollinearität zwischen den Prädiktoren hin. Nach der Rückwärtselimination bleiben als vorletzte Variable die Eigenrealisation ($\beta = 0,161$, $p = 0,212$) und als letzte Variable die Testleistung beim Bildkartentest ($\beta = -0,262$, $p = 0,045$) übrig, welche aber trotz Signifikanz nur 6,8 Prozent der Beurteilungsleistung von Teil E vorhersagen kann ($R^2 = 0,068$). Der Standardschätzfehler ist mit 1,027 relativ groß, weshalb die Berechnungen vorsichtig interpretiert werden müssen. Die Leistung bei der Bildsortieraufgabe kann jedenfalls als signifikanter Prädiktor für die synchrone Gesamtanalyse dienen. Überraschend ist jedoch, dass die Beurteilungsleistung um 0,149 Logit sinkt, wenn bei der Bildersortieraufgabe ein Punkt mehr erreicht wird. Außerdem kann wieder die Eigenrealisation – wenngleich nicht signifikant nachgewiesen – als positiver Prädiktor für die synchrone Gesamtanalyse dienen, denn wenn der Personenwert der Eigenrealisation um ein Logit steigt, so erhöht sich auch die Beurteilungsleistung von Teil E um 0,448 Logit.

5.1.6.6 *Bildung gemeinsamer Beurteilungsindices*

Wenn man die Teile A und B aufgrund ihrer relativ einfachen Aufgabenstellung zu einem gemeinsamen Index (BT_AB) zusammenfasst, so ergeben sich zwei signifikante Korrelationen mit dem biomechanischen Fachwissen ($r = 0,259$, $p = 0,047$) und der Eigenrealisation ($r = -0,325$, $p = 0,012$). Eine Regressionsanalyse bestätigt, dass die beiden Prädiktoren mit $p = 0,009$ ($F = 5,186$) ein signifikantes Modell darstellen und 15,6 Prozent der Beurteilungsleistung der Teile A und B erklären ($R^2 = 0,156$). Die Zunahme des biomechanischen Fachwissens um ein Logit lässt den entsprechenden Personenparameter um 0,191 Logit steigen ($\beta = 0,253$, $p = 0,045$), während die Verbesserung der Eigenrealisationsqualität um ein Logit eine Verringerung des Personenparameters um 0,504 Logit bewirkt ($\beta = -0,337$, $p = 0,009$). Die Kollinearität zwischen den Prädiktoren ist gering (Konditionsindex von 2,627). Die Effektstärke erweist sich mit $f^2 = 0,185$ als mittlerer Effekt.

Da die Teile C, D und E wesentlich komplexere Beurteilungsfähigkeiten erfordern als die Teile A und B, werden auch diese zu einem gemeinsamen Index (BT_CDE) zusammengefasst. Anhand der Korrelationsmatrix lassen sich die Eigenrealisation mit $r = 0,281$ ($p = 0,031$) und die Testleistung des TKBV-Teil1 mit $r = -0,259$ ($p = 0,047$) als signifikant korrelierende Variablen identifizieren. Wird die Wirkrichtung mittels Regressionsanalyse überprüft, so ist das Zwei-Prädiktorenmodell signifikant ($F = 3,857$, $p = 0,027$). Es erklärt 12,1 Prozent des gemeinsamen Beurteilungsindex BT_CDE, wobei die Erhöhung der Qualität der Eigenrealisation um ein Logit zu einer Erhöhung des Index um 0,328 Logit und die Erhöhung der Leistung des TKBV-Teil 1 um ein Logit zu einer Erniedrigung des Index um 0,104 Logit führt. Gemäß den standardisierten Beta-Koeffizienten sind beide Prädiktoren gleichwertig, wobei für die Eigenrealisation mit $\beta = 0,254$ ($p = 0,048$) im Gegensatz zum TKBV-Teil 1 mit $\beta = -0,223$ ($p = 0,081$) Signifikanz behauptet werden kann. Die Kollinearität kann mit einem Konditionsindex von 4,922 als gering erachtet werden. Es liegt mit $f^2 = 0,138$ eine geringe Effektstärke vor. Wenn man TKBV-Teil 1 als Prädiktor eliminiert, trägt die Eigenrealisation allein nur noch 7,1 Prozent zur Vorhersage der gemeinsamen Testleistung der Teile C, D und E bei ($R^2 = 0,071$).

Es lassen sich also für die beiden Indices gewisse Tendenzen feststellen, die auch schon in den Einzelteilen bestätigt wurden: die Beurteilungsaufgaben A und B scheinen durch ein gutes Fachwissen leichter zu lösen zu sein, während eine hohe Qualität der Eigenrealisation eine Verschlechterung der Beurteilungsleistung bewirkt. Bei den komplexeren Beurteilungsaufgaben C, D und E hingegen ist die Eigenrealisation ein signifikanter Prädiktor für das Erreichen einer guten Beurteilungsleistung, während eine erfolgreiche Bewältigung von Teil 1 des TKBV den Beurteilungsindex BT_CDE tendenziell verschlechtert. Diese Ergebnisse müssen sinnvoll interpretiert werden, um daraus möglicherweise Rückschlüsse der untersuchten Einflussfaktoren auf die verschiedenen Beurteilungsaufgaben zu ziehen (siehe Kapitel 6).

5.2 Detaillierte Auswertung und Analyse der Testverfahren

Die Datenaufbereitung und -auswertung erfolgte wiederum nach einem möglichst standardisierten Verfahren, um messtechnische Fehler im Vorfeld zu vermeiden. Abgesehen von der Bestimmung der Eigenrealisation wurden alle anderen Variablen mittels eines von den Probandinnen und Probanden auszufüllenden Testbogens bestimmt. Die Daten wurden dabei nach einer festgelegten Auswertungsroutine erfasst, indem die ausgefüllten Fragebögen zunächst eingescannt und mittels der Software KLAUS in eine CSV-Rohdaten-Tabelle exportiert wurden. Anschließend wurden mit dem Tabellenkalkulationsprogramm Excel alle richtigerweise nicht angekreuzten Items umcodiert, weil hier der erzielte Rohwert 0 eine richtige und der erzielte Rohwert 1 eine falsche Antwort darstellte. Die vorliegende Datenmatrix enthielt somit in den Zeilen die 59 Testpersonen und in den Spalten die jeweiligen Items des entsprechenden Testverfahrens. Die Datenmatrix wurde im CSV-Format in die Statistiksoftware R importiert und mit den Anwendungspackages eRm (extended Raschmodell, entwickelt von Mair, Hatzinger, Reinhold, Maier, Marko und Rusch (2013)) und ltm (latent trait model, entwickelt von Rizopoulos (2006)) die Rasch-Konformität der Daten zunächst überprüft, ehe die Item- und Personenparameter auf der Grundlage des konditionalen Maximum-Likelihood-Prinzips iterativ geschätzt wurden.

5.2.1 Fachwissenstest

Durchgeführt wurde der Fachwissenstest mit allen 59 Probandinnen und Probanden im Zeitraum vom 25. April bis zum 24. Mai 2017. Bei der Gesamtorganisation wurde darauf geachtet, dass alle Versuchsteilnehmer mit dieser Wissensfeststellung starteten, bevor andere Testverfahren aufgenommen wurden. Vor allem der später folgende Beurteilungstest sollte nach dem Fachwissenstest stattfinden, da durch den Beurteilungstest auch ein gewisser Lerneffekt über die Bewegungsausführung des Schlagwurfs erfolgt, wodurch die eigentliche Wissensfeststellung verzerrt werden könnte.

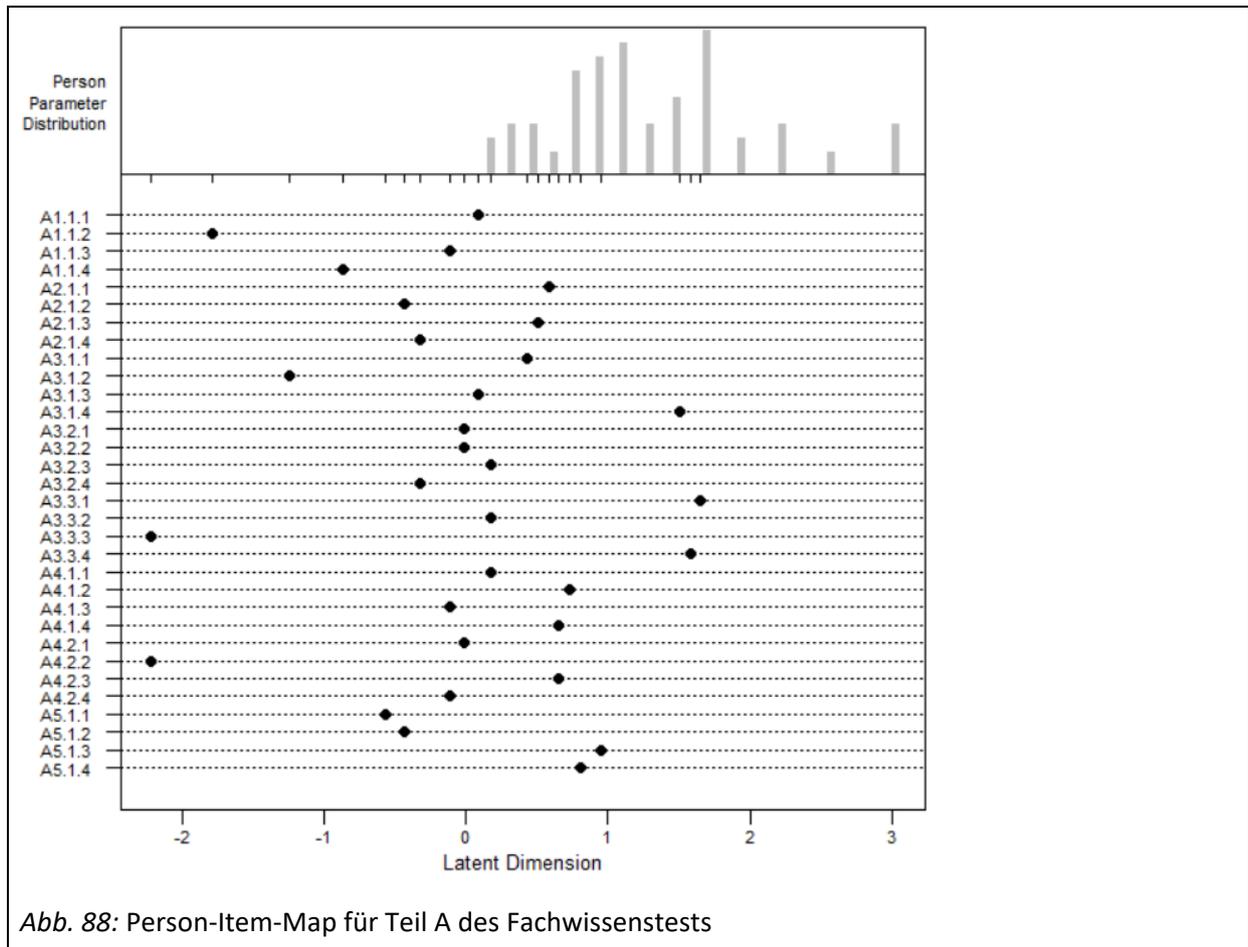
Eine detaillierte Darstellung der Ergebnisse des Fachwissenstests erfolgt in Kapitel 5.2.1.1, ehe in Kapitel 5.2.1.2 der statistische Nachweis erfolgt, dass die Anwendung des Rasch-Modells zur Berechnung der Personen- und Itemparameter seine Berechtigung hat und keine Eliminierung oder Reanalyse von Items vorgenommen werden muss. Außerdem wird durch die Prüfung zweier möglicher Auswertungsalternativen des Fachwissenstests (siehe Kapitel 5.2.1.3) festgestellt, dass der ermittelte Personenfähigkeitswert nicht vom statistischen Auswertungsverfahren abhängt, was für eine hohe Validität des Fachwissenstests spricht. Die Ergebnisse des Fehlerwissens und der visuellen Bewegungsvorstellung (siehe Kapitel 5.1.2.1) werden aufgrund der Erhebung über das freie Antwortformat bzw. des speziellen Auswertungsverfahrens der Bildersortieraufgabe (siehe Kapitel 4.2.4) lediglich auf deskriptive Weise dargestellt.

5.2.1.1 Bestimmung der Personen- und Itemparameter

Wenn man davon ausgeht, dass alle Items modellkonform sind (siehe Kapitel 5.2.1.2), so können die über das Rasch-Modell berechneten Personen- und Itemparameter als aussagekräftige Messwerte für das Fachwissen der Testpersonen betrachtet werden. Bei der Messung des Fachwissens wurde das räumlich-zeitliche Wissen (Teil A) mit 32 Items und das biomechanische Wissen (Teil B) mit 16 Items erfasst. Die beiden Teile wurden dabei zunächst getrennt behandelt, um bei späteren Untersuchungsschritten das Fachwissen differenzierter analysieren zu können.

Der durchschnittliche Personenparameter von Teil A liegt bei einem Wert von $\theta = 1,291$ ($SD = 0,669$) und lässt vermuten, dass die Items von der Mehrheit der Probandengruppe gut gelöst werden kann. Drei Probanden erreichten den Höchstwert von $\theta = 3,028$ (mit 30 von 32 richtigen Antworten). Den Negativwert von $\theta = 0,175$ (mit 17 von 32 erreichten Antworten) erzielten zwei Testpersonen. Die Separation Reliability beträgt 0,5199 (bei einer beobachteten Varianz von 0,7109 (Squared Standard Deviation) und einem Mean Square Measurement Error von 0,3413).

Anhand des Histogramms der Person-Item Map (siehe Abb. 88, Person Parameter Distribution) lässt sich erkennen, dass alle Testpersonen eine relativ hohe Personeneigenschaft ($\theta > 0$) zeigen, während die Personengruppe mit niedrigem Wert ($\theta < 0$) völlig fehlt. Ein sehr hoher Personenscore ($\theta > 2$) wird nur von wenigen Probanden erreicht. Erklärbar ist dies durch die Verteilung der Itemparameterschwierigkeiten (schwarze Punkte in Abb. 88) auf der latenten Dimension. Die Aufgabenschwierigkeit für Teil A liegt zwischen $\beta = -2,215$ (bei den Items A3.3.3 und A4.2.2) und $\beta = 1,654$ (beim Item A3.3.1), wobei ungefähr die Hälfte aller Items (mit $\beta < 0$) zu leicht für alle Probanden sind. Die andere Itemhälfte (mit $\beta > 0$) erscheint passend für einen großen Teil der Probanden. Jedoch fehlen Items für die Personengruppe mit hoher Personeneigenschaft (mit $\beta > 2$).



Der durchschnittliche Personenparameter für Teil B liegt bei einem Wert von $\theta = 0,443$ ($SD = 0,659$). Lediglich ein Proband erreichte den Höchstwert von $\theta = 2,217$ (mit 14 von 16 richtigen Antworten). Acht Personen erreichten einen Personenparameterwert von $\theta > 1$, wofür mindestens 12 Antworten richtig gelöst werden müssen. Den Negativwert von $\theta = -0,922$ (mit 5 von 16 erreichten Antworten)

erzielten zwei Testpersonen. Die Separation Reliability beträgt 0,2972 (bei einer beobachteten Varianz von 0,5639 (Squared Standard Deviation) und einem Mean Square Measurement Error von 0,3963).

Anhand des Histogramms der Person-Item-Map (siehe Abb. 89, Person Parameter Distribution) lässt sich erkennen, dass die meisten Testpersonen bei Teil B eine mittlere Personeneigenschaft ($0 < \theta < 1$) zeigen, während die Anzahl der Probanden mit niedrigem Personenwert ($\theta < 0$) und mit hohem Personenscore (mit $\theta > 1$) deutlich geringer ist. Somit ist die Verteilung der Personenparameter bei Teil B deutlich ausgeglichener als bei Teil A. Die Aufteilung der Itemparameterschwierigkeit auf der latenten Dimension zeigt, dass fünf Items (mit $\beta < -1$) zu leicht für alle Probanden sind. Die restlichen Items (mit $\beta > 0$) erscheinen passend für einen großen Teil der Probanden. Jedoch fehlen für die Personengruppe mit hoher Personeneigenschaft noch mehr schwierigere Items (mit $\beta > 1$). Die Aufgabenschwierigkeit für Teil B schwankt sehr stark zwischen den Extremwerten $\beta = -1,635$ (Item 6.2.2) und $\beta = 1,446$ (Item 6.2.1).

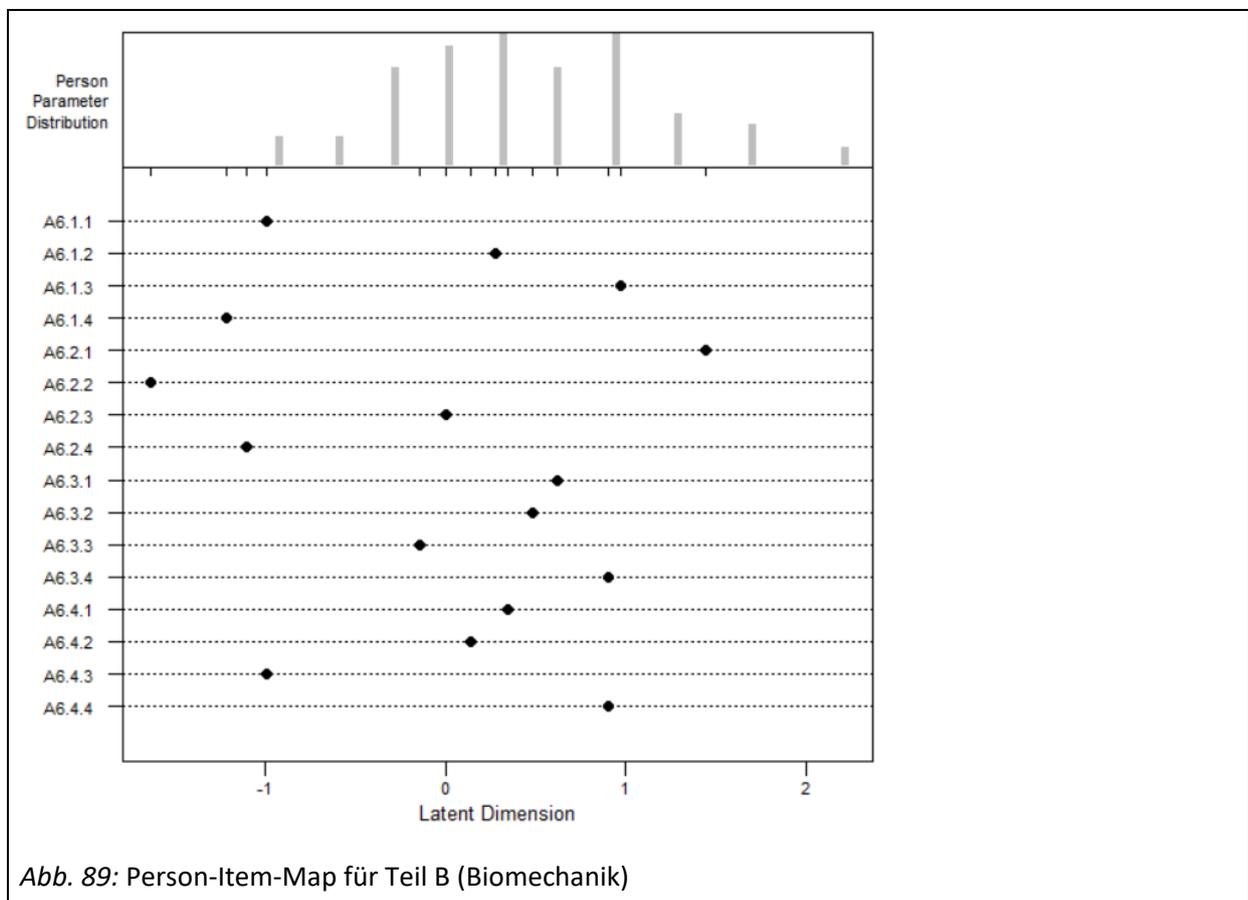


Abb. 89: Person-Item-Map für Teil B (Biomechanik)

Teil A ist also mit einem mittleren Personenwert von $\theta = 1,291$ für die Testpersonen leichter zu lösen als Teil B mit $\theta = 0,443$. Da es sich bei den Probandinnen und Probanden durchwegs um Sportstudierende handelt, die im Rahmen ihrer sportdidaktischen Ausbildung bereits einen praktischen und theoretischen Zugang zur Schlagwurfbewegung gewonnen haben, überrascht es nicht, dass die meisten Items zum räumlich-zeitlichen Verlauf der Schlagwurfbewegung für einen Großteil der Testgruppe zuverlässig gelöst werden konnten. Diese Aussage lässt sich statistisch dadurch begründen, dass allein 15 Items einen negativen Aufgabenschwierigkeitswert besitzen und von allen Testpersonen ($N = 59$ mit $\theta > 0$) mit einer Lösungswahrscheinlichkeit größer als 50 Prozent richtig beantwortet werden können. Selbst für das schwierigste Item (A3.3.4 mit $\beta = 1,559$) liegt bei 32 Prozent der Testpersonen

(N = 19) der Personenfähigkeitswert oberhalb der Aufgabenschwierigkeit, weshalb die Aufgabe für diese Personengruppe durchaus noch als unterfordernd angesehen werden kann. Bezüglich des räumlich-zeitlichen Bewegungsverlaufs ist demnach ein überdurchschnittlich hohes Wissen bei den Testpersonen verfügbar.

Wenn man auch den Inhalt der ursprünglich 32 Items berücksichtigt, so fällt auf, dass alle vier Items zur Ballrückführung während des Anlaufs (mit einer durchschnittlichen Aufgabenschwierigkeit von $\beta = -0,668$) am besten beantwortet wurden. Die Fragen zum Impulsschritt (mit $\beta = 0,082$) sowie zum Aufsetzen des Stemmbeins (mit $\beta = -0,040$) haben einen sehr ausgewogenen Schwierigkeitsgrad und wurden von den Testpersonen auch sehr zuverlässig richtig beantwortet. Deutlich schlechter ist jedoch die richtige Einschätzung der Aussagen über die Druckbeinbewegung in der Wurfauslage ($\beta = 0,196$) sowie über die räumlich-zeitliche Ausführung des Stemmschritts ($\beta = 0,299$) ausgefallen, obwohl bei zwei Items ein überdurchschnittliches Wissen bezüglich des Fußaufsatzes des Druckbeins ($\beta = -1,239$) sowie des zeitlichen Abstands zwischen dem Aufsetzen des Druck- und Stemmbeins ($\beta = -2,215$) festzustellen war. Das Wissen über die Abwurfbewegung ist als ambivalent zu betrachten, da Aussagen zur Körperverwindung schlecht beantwortet wurden ($\beta = 0,364$), während Details zur Ausführung der Wurfbewegung ($\beta = -0,418$) durchaus bekannt waren. Der Durchschnittswert letzteren Inhalts ist jedoch wenig aussagekräftig, weil die Aussage, dass der Ellbogen der Wurfhand zu Beginn der Wurfarmbeschleunigung vorseilt, fast allen Testpersonen (N = 57) bekannt war. Die Items über das Abfangen der Wurfbewegung offenbarten vor allem bezüglich der Stemm- und Druckbeinaktionen Wissenslücken und waren für die Testpersonen durchaus anspruchsvoll ($\beta = 0,187$).

Bei Teil B des Fachwissenstests besitzen sieben der 16 Items einen negativen Aufgabenschwierigkeitsparameter ($\beta < 0$). Im Vergleich zu Teil A sind immerhin 20 Prozent der Testpersonen (N = 12), für die ein negativer Personenparameterwert bestimmt wurde, mit den neun anderen biomechanischen Aussagen (mit $\beta > 0$) so stark gefordert, dass die Lösungswahrscheinlichkeit unter 50 Prozent liegt. Das schwierigste Item (A6.2.1 mit $\beta = 1,431$) kann lediglich von vier Testpersonen mit einer Lösungswahrscheinlichkeit von über 50 Prozent richtig beantwortet werden. Insgesamt variieren die Aufgabenschwierigkeiten aber so ausgewogen, dass abgesehen von den wenigen sehr fähigen Personen für jedes Wissensniveau passende Items enthalten sind. Die Itemschwierigkeiten passen bei Teil B also besser zum Fachwissensniveau der Testgruppe als bei Teil A, bei dem immerhin neun Testpersonen selbst das schwierigste Item mit einer Irrtumswahrscheinlichkeit unter 50 Prozent lösen können, während gleiches bei Teil B nur vier Testpersonen gelingt. Die Hälfte der Items von Teil A werden von der gesamten Probandengruppe tendenziell eher richtig als falsch eingeschätzt, während dies bei Teil B nur für ein Drittel der Items behauptet werden kann.

Auch hier ist ein Blick auf die inhaltliche Beantwortung der Items wieder aufschlussreich, um das Testergebnis richtig einordnen zu können. Die biomechanischen Fragen zu den letzten drei Anlaufschritten (mit einer durchschnittlichen Aufgabenschwierigkeit von $\beta = -0,253$) konnten von der Testgruppe gut gelöst werden, was vor allem an zwei sehr leicht zu beantwortenden Items über die Geschwindigkeitsgestaltung des Impulsschritts ($\beta = -0,966$) sowie die Funktion des Gegenarms ($\beta = -1,318$) liegt. Zugleich zeigt sich, dass ein erweitertes biomechanisches Verständnis über das schnelle Setzen des Stemmbeins bei den meisten Testpersonen nicht vorhanden ist (siehe Item A6.1.3 mit $\beta = 0,953$), obwohl ein ausgeprägtes Wissen speziell über den zeitlichen Abstand zwischen Druckbein- und Stemmbeinaufsatz vorhanden ist (siehe Item A3.3.3 mit $\beta = -2,215$). Die Beantwortung der Items über die Bewegung des Wurfarmes haben zwar eine ausgewogene durchschnittliche Aufgabenschwierigkeit ($\beta = 0,107$), streuen aber sehr stark. Die Kopplung der Teilimpulse von Ellbogen und Wurfhand ist 49 Testpersonen

bekannt (siehe Item A6.2.4 mit $\beta = -1,075$), während das Bewusstsein für das Abbremsen des Ellbogens am Ende der Wurfbewegung bei 44 Probandinnen und Probanden fehlt (siehe Item A6.2.1 mit $\beta = 1,431$). Die Testvignette zur Körperverwindung und Bogenspannung zeigt mit $\beta = 0,473$ die größten Wissenslücken im biomechanischen Testteil, wobei alle vier Items in einem relativ homogenen, mittleren Schwierigkeitsbereich liegen ($-0,109 < \beta < 0,880$). Ähnliches gilt für das Abfangen der Wurfbewegung mit einer durchschnittlichen Aufgabenschwierigkeit von $\beta = 0,459$ in einem Intervallbereich von $0,178 < \beta < 0,880$. Beide Aufgabenkomplexe sind für einen Großteil der Testgruppe fordernd, aber nicht unterfordernd. Allenfalls der Sachverhalt der Kopplung von Schulter und Wurfarm sowie die Funktion des Stemmbeins für das Abfangen des Körpers nach der Wurfbewegung erfordern für eine hohe Lösungswahrscheinlichkeit auch ein überdurchschnittliches biomechanisches Wissen (in Form eines hohen Personenfähigkeitswertes).

Wenn man die Ergebnisse von Teil A und B in eine Matrix nimmt, um einen Gesamtscore für die beiden Teile zu ermitteln, so ergibt sich ein durchschnittlicher Personenparameter von $\theta = 0,986$ ($SD = 0,554$). Der BIC-Index zeigt an, dass die Daten besser zum Rasch-Modell passen ($BIC = 3285,42$) als zum 2PL-Modell ($BIC = 3395,51$). Dadurch, dass Teil A mehr Items enthält als Teil B, die zugleich noch zu leicht für die Testgruppe sind, bleibt der durchschnittliche Personenparameter noch vergleichsweise hoch, der Gesamttest ist also für die Probandengruppe gut zu beantworten. Die Verteilung der Personenwerte zeigt, dass die Mehrzahl der Personen ($N = 45$) einen Personenparameter zwischen $0,421 < \theta < 1,587$ erreicht (siehe Appendix, Abb. 120). Lediglich ein Proband erreichte den Höchstwert von $\theta = 2,754$ (mit 44 von 48 richtigen Antworten), zwei Personen erzielten mit 42 richtigen Antworten einen Personenparameter von $\theta = 2,267$. Den Negativwert von $\theta = -0,183$ (mit 22 von 48 erreichten Antworten) erzielte lediglich eine Testperson. Die oben besprochenen Testvignetten und Einzelitems haben, wenn man sie als einheitliches Testinstrument auswertet, eine Aufgabenschwierigkeit zwischen den Extremwerten $\beta = -2,456$ (bei Items A3.3.3 und A4.2.2) und $\beta = 1,991$ (bei Item A6.2.1). 26 Items weisen einen negativen Aufgabenschwierigkeitsgrad auf, 21 davon haben einen geringeren Schwierigkeitsparameter als die geringste auftretende Ausprägung des Personenparameters ($\beta < -0,183$).

5.2.1.2 Prüfung der Modellauswahl

Die zur Prüfung der Modellauswahl angewendeten Verfahren bestätigen, dass für die beiden Teile A (Räumlich-zeitliches Wissen) und B (Biomechanisches Wissen) des Fachwissenstests keine Eliminierung und somit auch keine Reanalyse von Items vorgenommen werden muss, sondern die vorliegenden Daten sehr gut dem Rasch-Modell entsprechen, weshalb die Ermittlung der Personenparameter (siehe Kapitel 5.2.1.1) mittels Rasch-Modell zur Messung des Fachwissens begründet angewendet werden kann. Die Ergebnisse des BIC-Index des Likelihood-Quotienten-Tests (siehe Appendix, Tab. 59) bestätigen, dass das Modell mit der Annahme der identischen Ladungen (Rasch-Modell) sowohl für Teil A (Räumlich-zeitliches Wissen) als auch für Teil B (Biomechanisches Wissen) die Daten besser erklärt als das 2PL-Modell mit frei geschätzten Ladungen. Auch die im Folgenden beschriebenen quasi-exakten Testverfahren weisen die Raschkonformität der beiden Testteile nach.

Für beide Testteile darf außerdem begründet angenommen werden, dass kein Differential Item Functioning (DIF) vorliegt, weil die Items sowohl für die Gruppe der männlichen bzw. weiblichen Testper-

sonen sowie der Personen mit niedrigem bzw. hohem Personenwert ähnlich schwierig sind. Nachgewiesen wurde dies für die Teile A und B mittels der **globalen Teststatistik T_{10}** für das externe Splitkriterium Geschlecht (weiblich: $N = 26$, männlich: $N = 34$) mit Werten von $p = 0,296$ (Teil A) und $p = 0,309$ (Teil B) sowie für das interne Splitkriterium Personenscore-Median (Gruppe mit niedriger Personenfähigkeit: $N = 27$, Gruppe mit hoher Personenfähigkeit: $N = 32$) mit $p = 0,531$ (Teil A) und $p = 0,399$ (Teil B). Aufgrund der zwei möglichen Teilungskriterien ergibt sich nach Bonferroni-Korrektur (bei einem globalen α von 10 %) ein $\alpha^* = 0,1/2 = 0,05$. Für beide Teilungskriterien kann daher die Modellgültigkeit angenommen werden, weil beide p -Werte über dem α^* -Level liegen. Bestätigt wird diese Annahme vom parametrischen Andersen LR-Test, der für die Splitkriterien Median ($p = 0,513$) und Geschlecht ($p = 0,099$) die Nullhypothese bestätigt, dass das Rasch-Modell gilt (keine Signifikanz).

Überprüft man mittels der lokalen Teststatistik T_4 jedes Item explorativ auf Modellkonformität, indem für jedes Item bei jedem Teilungskriterium eine eigene Teststatistik berechnet wird, so kann festgestellt werden, dass sowohl für Teil A als auch für Teil B keine Items identifiziert werden können, die für die untere Fähigkeitsgruppe eindeutig leichter sind als unter dem Rasch-Modell erwartet. Item A3.1.4 (Druckbein drückt nach dem Aufsetzen nach vorne gegen das Stemmbein) ist das einzige, das für die untere Gruppe schwieriger als unter dem Rasch-Modell erwartet identifiziert wurde, und spricht aufgrund einer zu hohen Diskrimination gegen die Geltung des Rasch-Modells. Praktisch ist dieses Item jedoch einsetzbar, da es im statistischen Sinne zu gut zwischen Personen mit niedriger und höherer Fähigkeit diskriminiert.

Die detailliertere Betrachtung der Itemteststatistiken zeigt, dass bei Teil A für die Subgruppe Geschlecht nur die Items A2.1.3 ($p = 0$) und A3.2.2 ($p = 0,049$) bedeutsam vom Rasch-Modell abweichen ($\alpha^* = 0,05$), die restlichen 30 Items bestätigen die global berechnete Modellkonformität (siehe Appendix, Tab. 60). Die Aussage, dass der Impulsschritt durch ein ziehendes Aufsetzen und einen kräftigen Abdruck im drittletzten Bodenkontakt eingeleitet wird (Item A2.1.3), konnte von 30 männlichen Teilnehmern richtig gelöst werden, jedoch nur von zehn weiblichen Testpersonen. Item A3.2.2 (Das Stemmbein wird auf dem Fußballen aufgesetzt) wird von 29 männlichen und 16 weiblichen Testpersonen richtiger Weise als falsch identifiziert. Es kann jedoch kein inhaltlicher Grund gefunden werden, warum in beiden Fällen das entsprechende Item für Männer signifikant leichter zu beantworten ist als für Frauen, weshalb die Items im Test bleiben können. Von den vier Items mit $p > 0,9$ (A1.1.1, A2.1.1, A3.2.1 und A5.1.4) zeigen die Items A2.1.1 (Räumliche Ausführung des Impulsschritts) mit $p = 0,033$ und A5.1.4 (Druckbeinkontakt im Moment des Ballabwurfs) mit $p = 0,035$ bei der Alternativprüfung (R-Befehl alternative = "low") signifikante p -Werte, wodurch sich die Annahme bestätigt, dass die zwei Items für die Frauen zu leicht bzw. für die Männer zu schwer sind. Der graphische Modelltest (siehe Abb. 90) zeigt ebenfalls, dass bei Teil A die Items A2.1.1, A2.1.3 und A3.2.2 für das Splitkriterium Geschlecht nicht oder nur sehr knapp innerhalb des 95 %-Konfidenz-Bandes liegen und somit nicht uneingeschränkt modellkonform sind. Für die 16 Items des Teils B des Fachwissenstests kann mittels der Teststatistik T_4 die global berechnete Modellkonformität auf Itemebene ebenfalls bestätigt werden, denn alle Items erfüllen für das Teilungskriterium Geschlecht die Modellkonformität (siehe Appendix, Tab. 61). Die drei Items mit $p > 0,9$ (A6.1.1, A6.1.2 und A6.3.2) erweisen sich bei der Alternativprüfung als unauffällig ($p = 0,103$; $p = 0,122$; $p = 0,216$), wodurch sich die Annahme nicht bestätigt, dass die Items für die Frauen zu leicht bzw. für die Männer zu schwer sind.

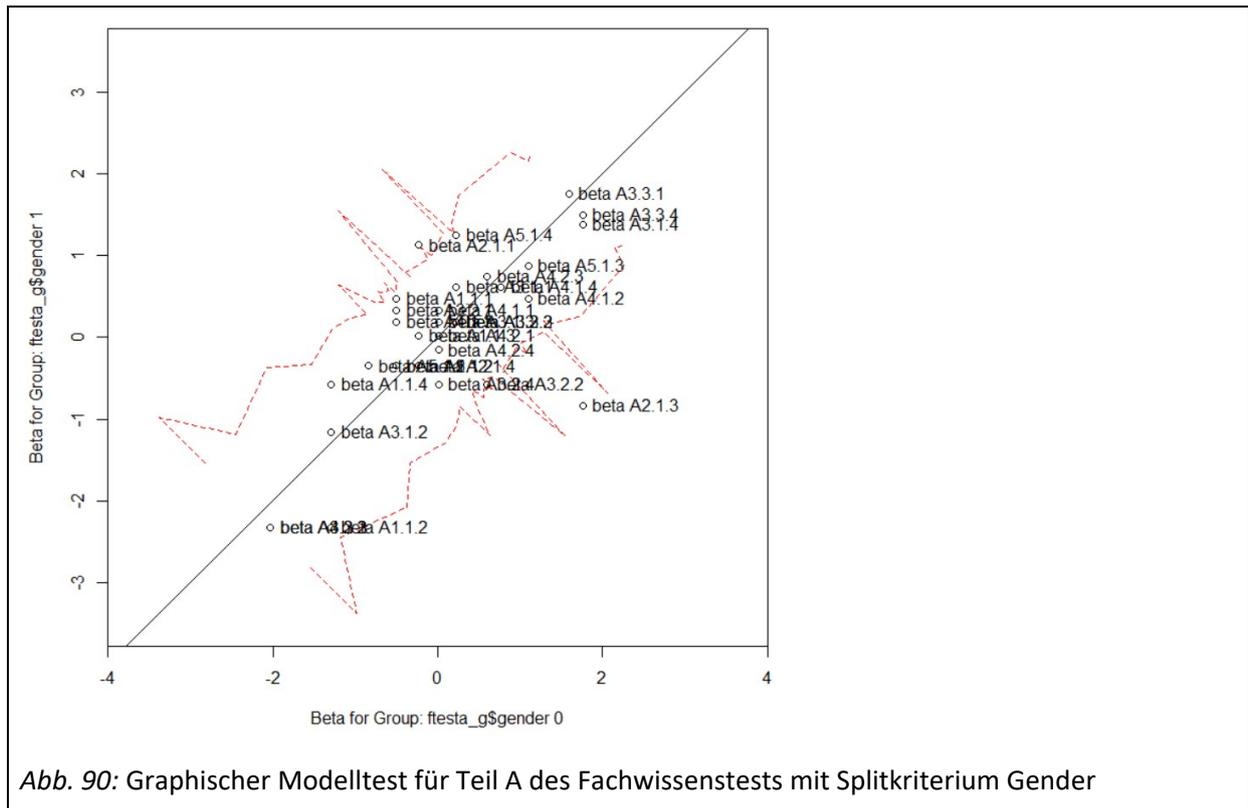


Abb. 90: Graphischer Modelltest für Teil A des Fachwissenstests mit Splitkriterium Gender

Wird T_4 mittels des Teilungskriteriums Mean berechnet, so ist für Teil A lediglich das Item A5.1.1 (Oberkörperrücklage im Abwurf) mit $p = 0,051$ knapp nicht-signifikant und zeigt die Tendenz, für die untere Fähigkeitsgruppe als zu leicht zu gelten. Von den acht auffälligen Items mit $p > 0,9$ (A3.1.2, A3.1.4, A3.3.1, A3.3.2, A3.3.3, A4.2.2, A4.2.4 und A5.1.3) hat sich hingegen lediglich Item A3.1.4 (Druckbein drückt nach dem Aufsetzen nach vorne gegen das Stemmbein) bei der Alternativprüfung (R-Befehl alternative = "low") mit $p = 0,038$ für die untere Fähigkeitsgruppe als bedeutsam schwieriger als unter dem Rasch-Modell erwartet erwiesen. Dies bestätigt auf parametrische Weise auch der Wald-Test, bei dem die Items A3.1.4 ($p = 0,048$) und A5.1.1 ($p = 0,040$) besonders niedrige p-Werte zeigen und die Auffälligkeit dieser beiden Items mittels der quasi-exakten Testverfahren bestätigen. In Teil B erfüllt für die untere Fähigkeitsgruppe lediglich das Item A6.2.2 (Handgelenksstreckung übt den letzten Beschleunigungsimpuls aus) mit $p = 0,059$ knapp die Signifikanzgrenze, muss aber noch nicht als zu leicht bezeichnet werden. Von den vier auffälligen Items mit $p > 0,9$ bestätigt die Alternativprüfung keines als bedeutsam schwieriger als unter dem Rasch-Modell erwartet (siehe Appendix, Tab. 61).

Zur Überprüfung der Itemhomogenität wird für Teil A die **Teststatistik T_{md}** (ähnlich dem Martin-Löf-Test) mit dem Splitkriterium Aufgabenschwierigkeit gemäß der Anzahl der ‚Löser‘ in eine schwierig zu lösende Itemgruppe ($idx\ 1 =$ Items 1, 5, 7, 9, 11, 12, 15, 17, 18, 20, 21, 22, 24, 27, 31, 32) und eine leicht zu lösende Itemgruppe ($idx\ 2 =$ 2, 3, 4, 6, 8, 10, 13, 14, 16, 19, 23, 25, 26, 28, 29, 30) berechnet. Das α^* beträgt wiederum $0,1/2 = 0,05$, da zwei Teilungskriterien zur Anwendung kommen. Die beobachtete Korrelation beträgt 0,334, der nicht-signifikante p-Wert von 0,61 deutet darauf hin, dass die Itemsgruppen als homogen eingestuft werden können. Mit $p = 0,287$ und einer beobachteten Korrelation von $r = 0,068$ deutet die Teststatistik T_{md} auch für Teil B darauf hin, dass die Itemsgruppen als homogen eingestuft werden können. Die Items wurden dazu mit dem Splitkriterium Aufgabenschwierigkeit gemäß der Anzahl der ‚Löser‘ in eine schwierig zu lösende Itemgruppe ($idx\ 1 =$ Items 2,3,5,9,10,12,13,16) und eine leicht zu lösende Itemgruppe ($idx\ 2 =$ 1,4,6,7,8,11,14,15) eingeteilt.

Der signifikante Wert $p = 0,002$ der globalen **Teststatistik** T_{11} zeigt an, dass für Teil A des Fachwissenstests die lokale stochastische Unabhängigkeit sowie die Homogenität der Items in Frage gestellt werden muss. Für das **Testverfahren** T_1 beträgt allerdings nach der Bonferroni-Korrektur bei 496 Itemkombinationen der Wert für $\alpha^* = 0,1/496 = 0,0002$, so dass lediglich für das Itempaar A3.1.2-A3.2.1 ein signifikanter Wert ($p < 0,0002$) festgestellt werden kann, der auf eine gegenseitige Abhängigkeit hindeutet. Bei beiden Items geht es um die Fußspitze des Druckbeins (A3.1.2) bzw. des Stemmbeins (A3.2.1) in der Wurfauslage, wobei lediglich acht Testpersonen die beiden Aussagen unterschiedlich einschätzen. Das Antwortmuster der Itemkombination A3.1.2-A3.2.1 von {11} tritt bei 45 Testpersonen auf, weshalb auch das Testverfahren T_{1m} einen signifikanten Wert ergibt. Gleiche Auffälligkeit gilt für das Itempaar A3.1.2-A3.2.3 (Das Stemmbein wird gestreckt aufgesetzt) mit 43 {11}-Antwortmustern. Trotz dieser Auffälligkeiten muss das Item A3.1.2 nicht aus dem Test entfernt werden, weil es keine inhaltliche Begründung für die Abhängigkeit bzw. hohe Lösungswahrscheinlichkeit mit den beiden genannten Items gibt. Vielmehr dürfte die Auffälligkeit daran liegen, dass die gefragten Inhalte der drei Items für die Testpersonen generell leicht zu lösen sind.

Der knapp nicht-signifikante Wert $p = 0,059$ der globalen **Teststatistik** T_{11} deutet darauf hin, dass für Teil B des Fachwissenstests die lokale stochastische Unabhängigkeit sowie die Homogenität der Items generell bestätigt werden kann. Über das **Testverfahren** T_1 kann von den insgesamt 120 Itemkombinationen ($\alpha^* = 0,1/120 = 0,0008$) für kein Itempaar ein signifikanter Wert ($p < 0,0008$) festgestellt werden. Die **Teststatistik** T_{1m} weist auf einem Signifikanzniveau unter 0,0008 ebenfalls keine Itemkombinationen aus, so dass Eindimensionalität der 16 Items angenommen werden kann. Auch ein Lerneffekt durch das sukzessive Beantworten der 16 Items kann mit dem **Testverfahren** T_{1l} widerlegt werden, weil es keine signifikanten Itempaare gibt, die ein erhöhtes Auftreten des Antwortmusters {11} nachweisen können. Somit zeigen sich alle Items von Teil B des Fachwissenstests mit dem Rasch-Modell konform.

In der **Teststatistik** T_{pbis} liegen alle p-Werte der 32 Items von Teil A und der 16 Items von Teil B über der korrigierten Signifikanzgrenze, was darauf hindeutet, dass alle Items zu der restlichen Skala der anderen Items passen und eine zufriedenstellende Trennschärfe aufweisen. In Teil A ($\alpha^* = 0,1/32 = 0,003$) liegen die Items A3.1.3 ($p = 0,010$) und A5.1.1 ($p = 0,014$) relativ nahe am korrigierten Alpha-Wert. Die Trennschärfe für Teil B ($\alpha^* = 0,1/16 = 0,006$) ist selbst für das Item A6.4.3 mit dem geringsten Wert ($p = 0,017$) noch erfüllt.

Die Modellannahme bestätigt sich auch über die parametrisch berechneten Infit-MNSQ-Werte. Für die 32 Einzelitems des Fachwissenstests Teil A (RZ-Wissen) weist der Itemparameter Infit-MNSQ einen Wertebereich zwischen 0,888 (Item A3.1.2) und 1,137 (Item A3.1.3) auf und bestätigt somit, dass alle Items in einem als sehr Rasch-konform zu bezeichnenden Bereich von 0,8 bis 1,2 liegen und somit ein absolut zufriedenstellender Fit der Items vorliegt. Mit Infit-t-Werten zwischen -1,36 (Item A3.1.4) und 0,95 (Item A3.1.3) sowie Outfit-t-Werten zwischen -1,21 (Item A3.1.4) und 1,65 (Item A2.1.2) wird der Grenzbereich von $-1,96 < t < 1,96$ in beiden Fällen deutlich erfüllt, so dass kein Item gestrichen werden muss. Gleiches trifft auf Teil B zu, denn die Itemfitparameter bewegen sich für den Infit-MNSQ zwischen ,862 (Item 6.2.1) und 1,114 (Item 6.4.3). Auch hier erfüllen alle Infit-MNSQ-Werte das strenge Intervall zwischen 0,8 und 1,2 und deuten auf einen sehr guten Fit der Items für das Konstrukt des biomechanischen Fachwissens hin. Die t-Werte liegen für den Infit-MNSQ zwischen -1,59 (Item A6.4.2) und 0,69 (Item A6.3.3) sowie für den Outfit-MNSQ zwischen -1,62 (Item A6.4.2) und 0,98 (Item A6.4.3) und sprechen für eine hohe Modellkonformität aller Items.

5.2.1.3 Auswertungsalternativen

Es kann durchaus diskutiert werden, ob alternative Auswertungsmethoden zu einem anderen Ergebnis führen würden. Daher wird in einem letzten Schritt überprüft, ob sich dadurch die Zusammenhänge mit anderen Einflussfaktoren ändern. Es werden dabei die Personenparameter für den Gesamttest (Teile A und B zusammen) neu berechnet und miteinander verglichen, um durch eventuelle Abhängigkeiten zu zeigen, dass die Auswertungsmethoden keinen signifikanten Einfluss auf die ermittelte Personenfähigkeit haben. In späteren Korrelationsanalysen wird auch immer überprüft, ob sich durch alternative Datenauswertung möglicherweise andere Zusammenhänge vermuten lassen.

Die erste Auswertungsalternative besteht darin, eine gewisse Homogenisierung der Itemschwierigkeiten herzustellen. Wenn alle Items entfernt werden, deren Itemschwierigkeit kleiner ist als der geringste Personenparameterwert der Testgruppe ($\theta = -0,183$), so reduziert sich der Aufgabenpool auf 27 Items, die nun relativ gut zur Fähigkeitsausprägung der Testgruppe passen. Zugleich bleibt gewährleistet, dass aus jedem der Wissensbereiche mindestens eine Frage das Schwierigkeitskriterium erfüllt, so dass die inhaltliche Grundstruktur des Fachwissenstests erhalten bleibt. Anhand des Histogramms in Abb. 91 sieht man, dass es nur noch vereinzelte Testpersonen sind, für die alle Aufgaben tendenziell zu schwer oder zu leicht sind. Der durchschnittlich ermittelte Personenparameter reduziert sich (von ursprünglich $\theta = 0,986$) auf $\theta = 0,303$ ($SD = 0,627$). Den Höchstwert von $\theta = 2,212$ (24 von 27 richtigen Antworten) erreichte eine Testperson, der Tiefstwert von $\theta = -1,131$ (7 von 27 richtigen Antworten) wurde ebenfalls lediglich für eine Testperson berechnet. 20 Testpersonen erzielten einen negativen Personenparameter. Ausgenommen von Item A6.2.1 sind alle anderen Items ausgewogen verteilt im Intervallbereich von $-1 < \beta < 1$. Da 51 Testpersonen einen Personenwert von $\theta < 1$ aufweisen, liegen die Testitems für 86 Prozent der Probandinnen und Probanden in einem ihren Fähigkeiten entsprechenden oder fordernden Schwierigkeitsbereich. Der Gesamtscore aus der 27-Items-Analyse kann somit als aussagekräftig für das Fachwissen der Testpersonen über die Schlagwurfbewegung gelten, weil die verwendeten Items bezüglich ihrer Aufgabenschwierigkeit sehr gut zur gemessenen Personenfähigkeit (d. h. dem vorhandenen Fachwissen der Testpersonen) passen und für einen Großteil der Testgruppe anspruchsvolle, aber nicht unlösbare inhaltliche Anforderungen stellen.

Als zweite Alternative werden die Fragen des Fachwissenstests (Teile A und B) nach dem Multiple-True-False-Format (MTF) (in der speziellen Variante des Kprim-Verfahrens) ausgewertet, bei dem mehrere True-False-Aussagen zu einem einleitenden Aufgabenstamm präsentiert werden (Albanese, Kent & Whitney, 1979). Der Vorteil gegenüber einfachen True-False-Aufgaben liegt darin, dass eine MTF-Aufgabe nur kreditiert wird, wenn der Prüfling alle zugehörigen Aussagen korrekt bewertet, wodurch die Ratewahrscheinlichkeit effektiv gesenkt werden kann (Lindner, Strobel & Köller, 2015). Da für jeden übergeordneten Wissensinhalt (z. B. Wurfarmbewegung oder Aufsetzen des Stemmbeins) vier Items konstruiert und abgefragt wurden, kann die Auswertung derart erfolgen, dass zwei Rohpunktwerte vergeben werden, wenn alle Items richtig beantwortet wurden, und ein Rohpunktwert, wenn eines der vier Items falsch beantwortet wurde. In allen anderen Fällen gilt die Aufgabe als nicht gelöst (Rohpunktzahl 0). Dieses sehr strenge Auswertungsverfahren lässt insbesondere bei einem hohen Anforderungsniveau der Items eine differenziertere Bewertung der Teilnehmer zu. Durch diese Zusammenfassung der Items zu inhaltlich strukturierten Testvignetten kann überprüft werden, ob sich das Auswertungsverfahren aller Einzelitems von der KPRIM-Auswertung unterscheidet. Die KPRIM-Choice-Frage gehört nämlich zu den drei testtheoretisch am besten ausgewerteten Multiple-Choice-Fragetypen.

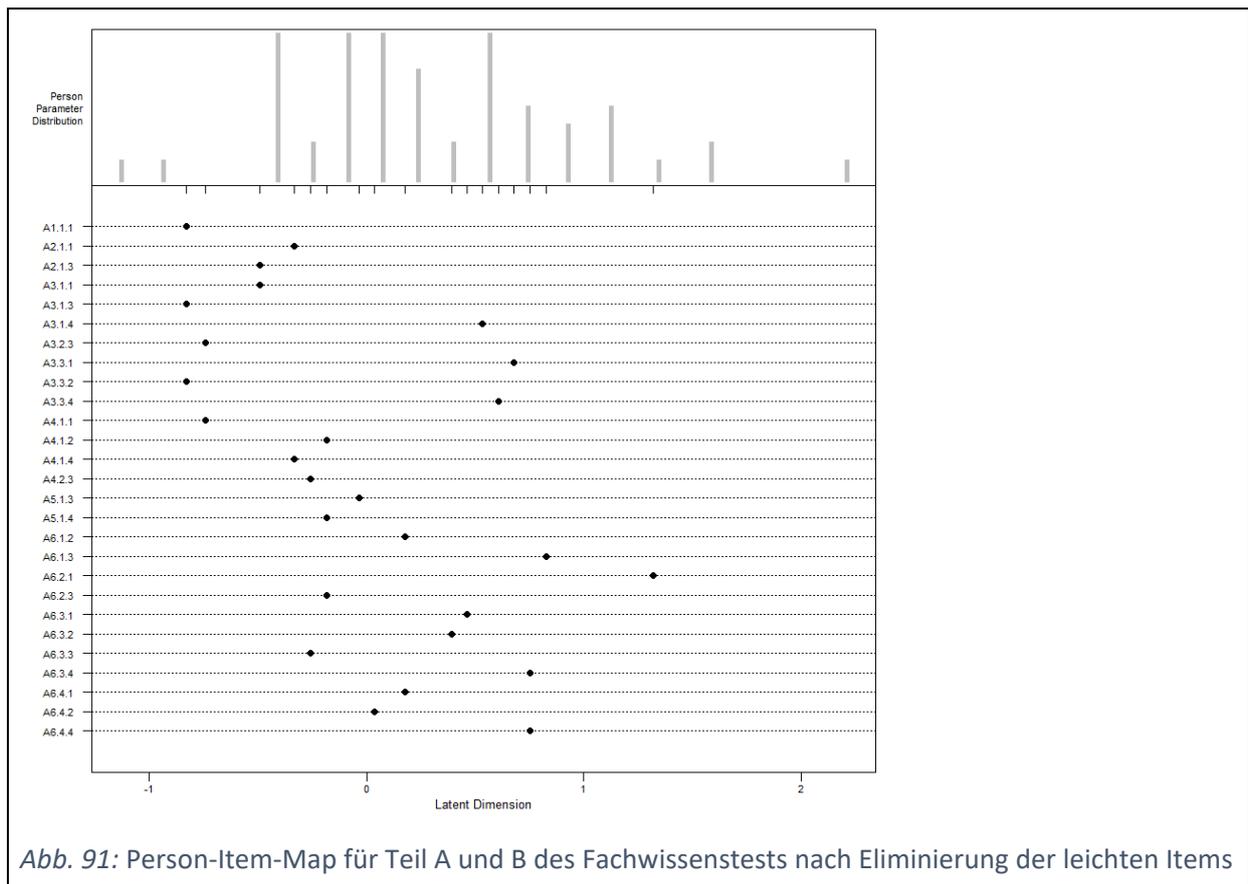


Abb. 91: Person-Item-Map für Teil A und B des Fachwissenstests nach Eliminierung der leichten Items

Die statistische Berechnung der Personenparameter erfolgt in diesem Fall nach dem ordinalen Rasch-Modell, weil nun keine dichotomen Daten mehr vorliegen, sondern drei Fähigkeitsstufen unterschieden werden. Durch dieses sehr strenge Auswertungsverfahren reduziert sich der durchschnittliche Personenparameter für Teil A (von ursprünglich $\theta = 1,291$) auf $\theta = 0,182$ ($SD = 0,780$), für Teil B (von ursprünglich $\theta = 0,443$) auf $\theta = -1,053$ ($SD = 1,116$) und für die gemeinsame Auswertung der Teile A und B (von ursprünglich $\theta = 0,986$) auf $\theta = -0,189$ ($SD = 0,676$). In Abb. 92 sieht man, dass sowohl die Personenwerte als auch die Itemschwierigkeiten sehr ausgewogen verteilt sind. Die Items A1.1 und A5.1 sind dahingehend problematisch, dass sie von Personen mit niedrigerem Personenfähigkeitswert besser beantwortet werden als von Personen mit hohem Fähigkeitswert. Die Extremwerte liegen aber bei beiden Items sehr eng beieinander, weshalb die Passung der Items nicht in Frage gestellt werden muss.

Wenn man nun die mit verschiedenen Auswertungsverfahren ermittelten Personenparameter vergleicht, indem man die entsprechenden Korrelationskoeffizienten ermittelt, so zeigt sich, dass der errechnete Personenparameter nur in sehr geringem Maß von den Auswertungsmethoden abhängt. Die entsprechenden Korrelationskoeffizienten zwischen den drei Verfahren für den Gesamttest (Teile A und B zusammen) sind mindestens größer als $r = 0,864$. Für Teil A beträgt der Korrelationskoeffizient zwischen der Auswertung aller Items und der KPRIM-Auswertung $r = 0,961$, für Teil B $r = 0,930$. Dies spricht für eine hohe Validität des Fachwissenstests, weil trotz unterschiedlicher Auswertungsverfahren eine sehr ähnliche Zuordnung der Personenparameterwerte zu den einzelnen Testpersonen vorliegt.

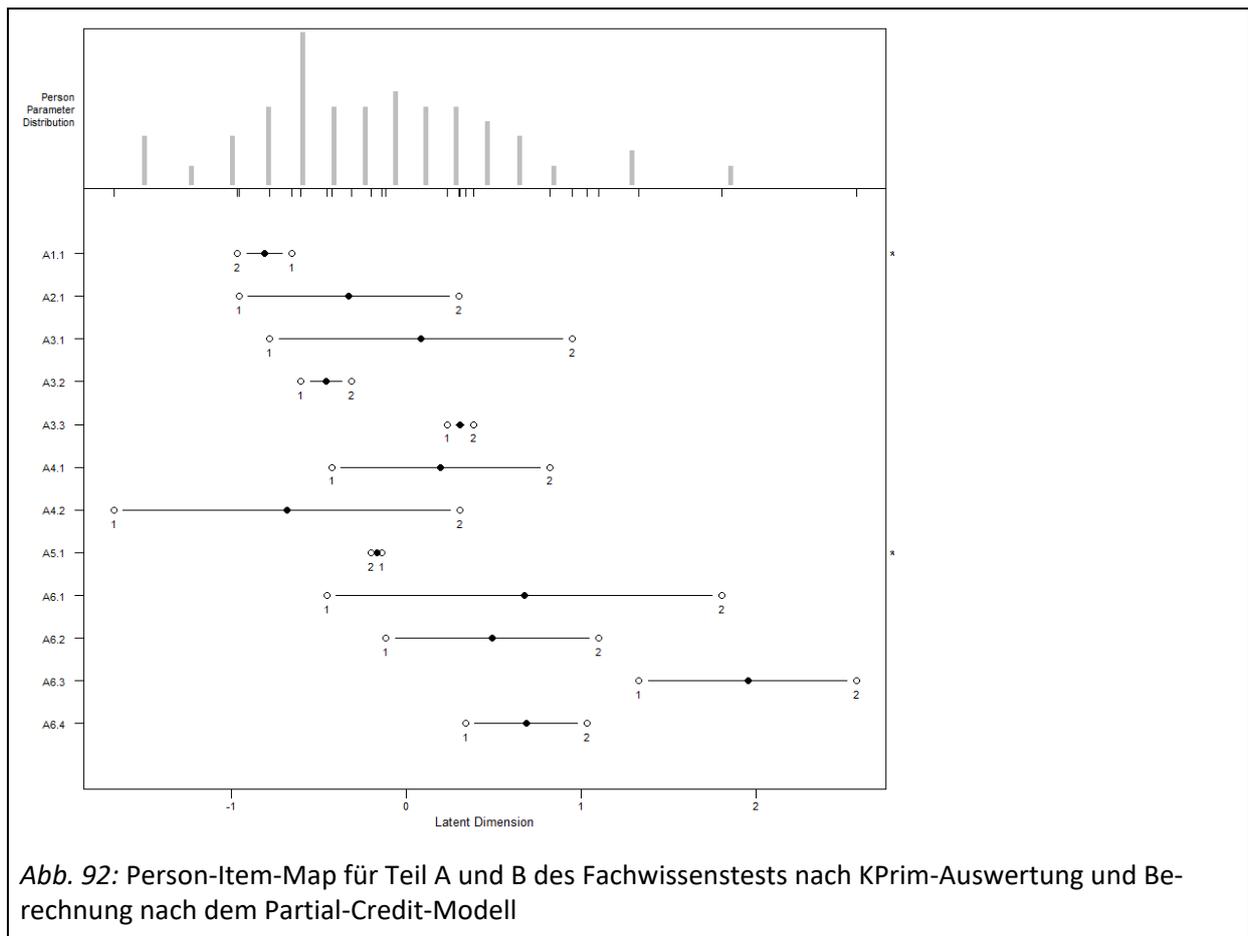


Abb. 92: Person-Item-Map für Teil A und B des Fachwissenstests nach KPrim-Auswertung und Berechnung nach dem Partial-Credit-Modell

5.2.2 Test zur Kontrolle der Bewegungsvorstellungsfähigkeit (TKBV)

Die Fähigkeit zum Ist-Sollwert-Vergleich der Testpersonen wird mit dem Test zur Kontrolle der Bewegungsvorstellung (TKBV) gemessen. Mittels der quasi-exakten Testverfahren kann bestätigt werden, dass die Daten der insgesamt 20 Aufgaben zum Rasch-Modell passen (siehe Kapitel 5.2.2.2). Die Vorstellungsaufgaben werden von der Probandengruppe dieser Studie sehr zuverlässig gelöst, was durch einen hohen mittleren Personenparameter zum Ausdruck kommt und auf eine gewisse Unterforderung der Testpersonen hindeutet. Dabei müssen die beiden Teile des TKBV differenzierter betrachtet werden, weil vor allem die Wiedererkennensanforderung von Teil 1 unter dem eigentlichen Leistungsvermögen der meisten Testpersonen lag, während das Einnehmen der korrekten Endposition nach dem gedanklichen Bewegungsvollzug deutlich mehr Schwierigkeiten bereitete. Für den neu konzipierten Bewegungsbeobachtungstest lässt sich eine schwache positive Korrelation mit den Personenwerten des TKBV nachweisen, was durchaus darauf hindeuten kann, dass zwar dem TKBV ähnliche, aber zugleich komplexere Fähigkeiten benötigt werden, um die schwierigere Aufgabenstellung des Bewegungsbeobachtungstests über die visuelle Informationsaufnahme zu bewältigen.

5.2.2.1 Item- und Personenparameter des TKBV

Die Bestimmung der Itemschwierigkeit (siehe Tab. 39) zeigt, dass acht Items einen negativen Wert haben und somit einer Mehrheit der Testpersonen leichter fallen als die anderen elf Items. Alle Items des Teils 1 außer A1 und A3 gehören zu dieser leichten Itemgruppe, was darauf schließen lässt, dass vor allem Teil 2 eine Differenzierung der Personenparameter bewirkt. Die mit Abstand schwersten Items sind A12 ($\beta = 1,583$) und A14 ($\beta = 1,660$). Item A6 wurde aus der Datenmatrix eliminiert, weil alle Testpersonen diese Aufgabe richtig beantworteten und die Frage somit für die statistische Auswertung unbrauchbar ist.

Tab. 39: Itemschwierigkeit für die 19 Items des TKBV

	beta	SD	CI low	CI high		beta	SD	CI low	CI high
A1	0.296	0.298	-0.288	0.881	A11	0.888	0.279	0.341	1.435
A2	-1.543	0.506	-2.534	-0.552	A12	1.583	0.278	1.037	2.128
A3	0.106	0.308	-0.498	0.710	A13	0.966	0.278	0.422	1.511
A4	-1.543	0.506	-2.534	-0.552	A14	1.660	0.279	1.113	2.208
A5	-0.454	0.351	-1.142	0.233	A15	0.297	0.298	-0.288	0.881
A7	-0.737	0.380	-1.482	0.009	A16	0.106	0.308	-0.498	0.710
A8	-0.590	0.364	-1.303	0.124	A17	0.005	0.314	-0.611	0.621
A9	-0.590	0.364	-1.303	0.124	A18	0.297	0.298	-0.288	0.881
A10	-0.590	0.364	-1.303	0.124	A19	0.297	0.298	-0.288	0.881
					A20	-0.454	0.351	-1.142	0.233

Der durchschnittliche Personenparameter liegt bei einem Wert von $\theta = 1,400$ ($SD = 1,033$) und lässt vermuten, dass die Aufgaben für die Probandengruppe eher leicht zu beantworten sind. Lediglich eine Testperson erreichte den Höchstwert von $\theta = 4,083$ (mit 19 von 19 richtigen Antworten). Elf weitere Probanden erzielten einen Personenparameterwert $\theta > 2$, wofür mindestens 17 Items richtig gelöst werden mussten. Der Negativwert von $\theta = -0,631$ (mit 7 von 19 erreichten Antworten) wurde lediglich für eine Testperson gemessen, für drei weitere Probanden wurde ein negativer Personenwert bestimmt (8 von 19 Items richtig beantwortet). Die Separation Reliability beträgt 0,6086 (bei einer beobachteten Varianz von 1,0862 (Squared Standard Deviation) und einem Mean Square Measurement Error von 0,4251).

Anhand des Histogramms der Person-Item Map (siehe Abb. 93, Person Parameter Distribution) lässt sich erkennen, dass die meisten Testpersonen eine mittlere Personeneigenschaft ($0 < \theta < 2$) zeigen, während die Personenanzahl mit niedrigem Wert ($\theta < 0$) und mit hohem Personenscore (mit $\theta > 2$) gering sind. Die Verteilung der Itemparameterschwierigkeit auf der latenten Dimension zeigt, dass die acht Items mit negativem Aufgabenschwierigkeitswert lediglich für die Gruppe mit niedrigem Personenwert passend sind und auch die neun Items mit $0 < \beta < 1$ für den Großteil der Testgruppe unterfordernd sind. Selbst die schwierigsten Items A12 und A14 können von der Personengruppe mit hoher Personeneigenschaft (insgesamt sind es 12 Testpersonen mit $\theta > 2$) zuverlässig gelöst werden, während die Items A2 und A4 wiederum so leicht sind, dass sie selbst von Testpersonen mit niedrigem Personenwert souverän gelöst werden können.

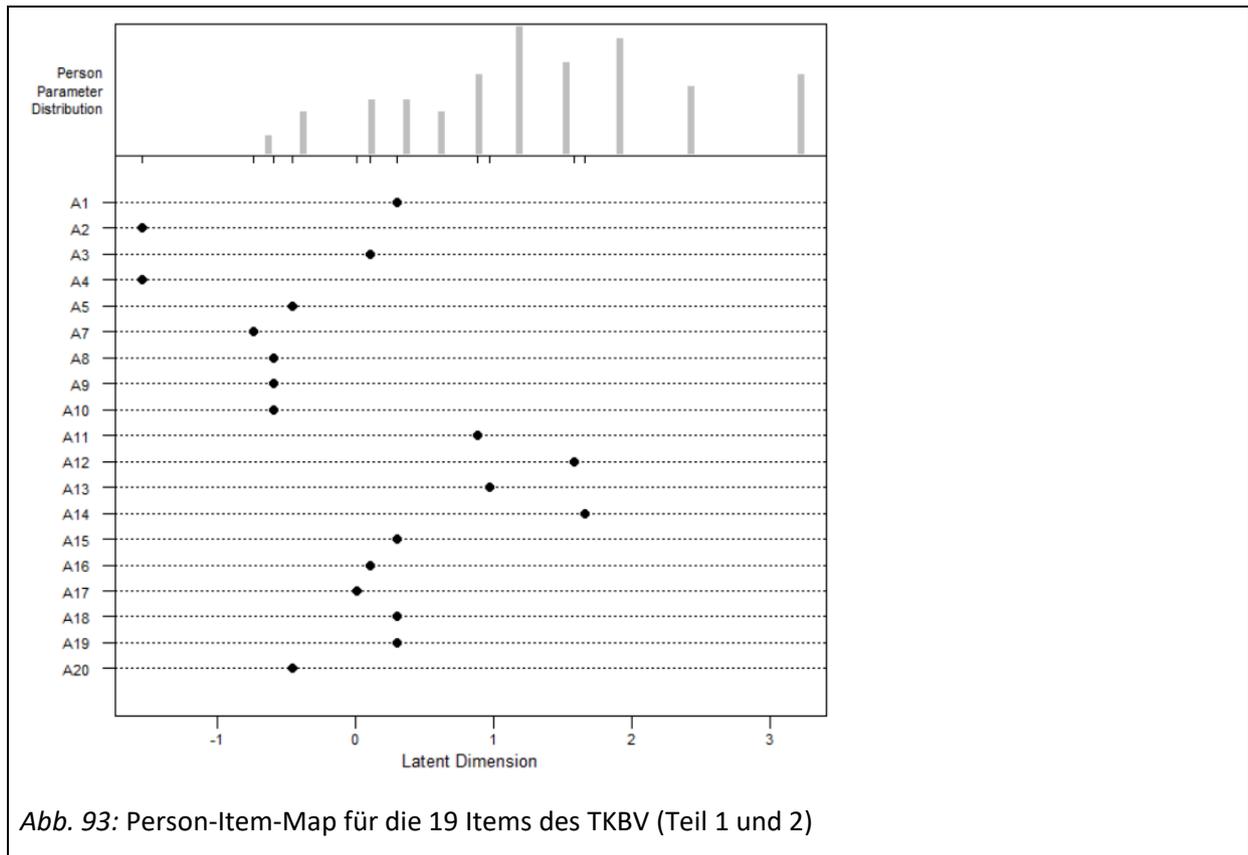


Abb. 93: Person-Item-Map für die 19 Items des TKBV (Teil 1 und 2)

5.2.2.2 Prüfung der Modellauswahl des TKBV

Die informationstheoretische Berechnung legt nahe, dass das Rasch-Modell (BIC-Index 1211,01) die Daten besser erklärt als das 2PL-Modell mit frei geschätzten Ladungen (BIC-Index 1246,63). Auch die Bootstrap-Goodness-of-Fit-Statistik mittels Pearson- χ^2 -Test (300 Bootstrap-Samples) bestätigt, dass die Hypothese der Modellanpassung nicht verworfen werden muss ($p = 0,246$). Die **globale Teststatistik T_{10}** liefert zunächst ein ambivalentes Ergebnis, wenn man nach Alpha-Korrektur eine Signifikanzgrenze von $p = 0,1/2 = 0,05$ annimmt. Bei Anwendung des Splitkriteriums Median/Mittelwert deutet ein signifikanter p-Wert ($p = 0,036$) auf die Invarianz der Subgruppen hin, während für das externe Splitkriterium Geschlecht ($p = 0,207$) mit einem nicht-signifikanten p-Wert die Modellgültigkeit bestätigt werden kann. Das Ergebnis wird vom parametrischen LR-Test mit Splitkriterium Geschlecht ($p = 0,373$) bestätigt.

Bei der Überprüfung dieser global berechneten Ergebnisse auf Itemebene weist die **Teststatistik T_4** (Signifikanzgrenze $\alpha^* = 0,1/20 = 0,005$) für das Splitkriterium Geschlecht alle Items als modellkonform aus, während für das Splitkriterium Median das Item A19 ($p = 0,005$) nicht zum Rasch-Modell passt, d. h. Aufgabe 19 des TKBV erscheint für die obere Fähigkeitsgruppe zu leicht. Von den vier Items mit $p > 0,9$ für das Splitkriterium Median (A3, A5, A9 und A11) sowie den fünf Items mit $p > 0,9$ für das Splitkriterium Geschlecht (A3, A4, A9, A11 und A20) zeigte bei der Gegenprüfung mit der jeweils anderen Subgruppe lediglich das Item A11 für das Splitkriterium Median mit $p = 0,005$ einen signifikanten p-Wert. Dies lässt den Rückschluss zu, dass für Testpersonen mit einem hohen TKBV-Score die Aufgabe 11 schwieriger und für Testpersonen mit einem niedrigen TKBV-Score leichter ist als unter dem Rasch-Modell erwartet. Da die Aufgaben des TKBV innerhalb der einzelnen Teile grundsätzlich sehr ähnlich

sind, sind die Besonderheiten der Items A19 und A11 nicht inhaltlich zu erklären und werden daher bei der Datenauswertung auch nicht besonders behandelt. Das Item A11 erscheint auch bei der parametrischen Überprüfung der Fit-Statistiken auffällig, da sowohl der Infit-MNSQ-Wert mit 1,266 als auch der Infit-t-Wert mit 2,28 über den bereits genannten Grenzwerten liegt.

Bei Überprüfung der Itemhomogenität mittels der **Teststatistik T_{md}** (schwierig zu lösende Itemgruppe ($idx\ 1 = 1,3,10-15,17,18$) und leicht zu lösende Itemgruppe ($idx\ 2 = \text{Items } 2,4-9,16,19$)) beträgt die beobachtete Korrelation $r = 0,413$, der nicht-signifikante p-Wert von 0,280 deutet darauf hin, dass die Itemsubgruppen (bei einem korrigierten Alpha-Wert von $0,1/2 = 0,05$) als homogen eingestuft werden können. Der signifikante Wert $p = 0,015$ der **Teststatistik T_{11}** ($n = 1000$, $burn_in = 500$, $step = 64$, $seed = 98765$) liegt unter der 0,1-Signifikanzgrenze und deutet darauf hin, dass für den TKBV entweder eine lokale stochastische Unabhängigkeit oder die Homogenität der Items angezweifelt werden kann. Über das **Testverfahren T_1** , für welches nach der Bonferroni-Korrektur bei 171 Itemkombinationen der Wert für $\alpha^* = 0,1/171 = 0,0006$ bestimmt wird, können alle Itempaare als unauffällig betrachtet werden. Somit gibt es keine bedeutsam voneinander abhängigen Itemkombinationen, alle Items sind lokal stochastisch unabhängig. Multidimensionalität kann ebenso für alle Itemkombinationen ausgeschlossen werden, denn die Teststatistik T_{1m} zeigt keine signifikanten p-Werte. Somit gehören alle Items derselben latenten Dimension an.

Als letztes betrachten wir noch die Trennschärfe der einzelnen Items mittels der **Teststatistik T_{pbis}** ($n = 1000$, $burn_in = 500$, $step = 64$). Abgesehen von Item A3 ($p < 0,001$) liegen alle p-Werte über der korrigierten Signifikanzgrenze von $\alpha^* = 0,1/19 = 0,005$. Dies deutet darauf hin, dass 18 Items zu der restlichen Skala der anderen Items passen, während Item A3 eine zu niedrige Trennschärfe aufweist. Inhaltlich lässt sich hierfür keine Erklärung finden, weil Aufgabe 3 genauso konzipiert ist wie alle anderen Aufgaben von Teil 1 des TKBV. Die Aufgabe wird daher bei der Testauswertung nicht eliminiert, auch um den TKBV möglichst originalgetreu zu verwenden.

5.2.2.3 Exkurs: Ergebnisse des Bewegungsbeobachtungstests

5.2.2.3.1 Modellprüfung des Bewegungsbeobachtungstests

Da der Test im Rahmen dieser Arbeit neu konstruiert wurde, ist die Modellprüfung des ermittelten Datensatzes besonders wichtig. Item B6 wird a priori aus der Datenmatrix eliminiert, weil bei der Beantwortung der Frage während des Tests Unklarheiten bei den Probanden auftauchten, die zu einer Verzerrung der Testergebnisse führten. Der BIC-Index legt nahe, dass das Rasch-Modell (BIC-Index 706,49) die Daten besser erklärt als das 2PL-Modell (BIC-Index 728,08). Die Modellgültigkeit kann auch mittels der **globalen Teststatistik T_{10}** ($n = 1000$, $burn_in = 500$, $step = 64$) bestätigt werden. Für die drei Splitkriterien Median (Gruppe 1: $N = 17$, Gruppe 2: $N = 42$), Mittelwert (Gruppe 1: $N = 29$, Gruppe 2: $N = 30$) und Geschlecht (Gruppe 1: $N = 25$, Gruppe 2: $N = 34$) wurde jeweils ein nicht-signifikanter p-Wert erreicht (Median mit $p = 0,780$, Mittelwert mit $p = 0,809$ und Geschlecht mit $p = 0,255$), was auf die globale Modellgültigkeit für die entsprechenden Subgruppen hindeutet. Auch auf Itemebene erweisen sich in der **Teststatistik T_4** alle Items sowohl für das Splitkriterium Median als auch für das Geschlecht als modellkonform und liegen über dem korrigierten Signifikanzniveau von $\alpha^* = 0,1/9 = 0,01$. Alle Items weichen somit nicht bedeutsam vom Rasch-Modell ab. Die Itemhomogenität wird mittels der **Teststatistik T_{md}** zwischen einer schwierig zu lösenden ($idx\ 1 = 1-5$) und einer leicht zu lösenden Itemgruppe ($idx\ 2 = \text{Items } 6-9$) bestätigt, weil ein nicht-signifikanter p-Wert von 0,570 für die beobachtete Korrelation ($r = 0,066$) vorliegt.

Der nicht-signifikante Wert von $p = 0,479$ der globalen **Teststatistik** T_{11} deutet darauf hin, dass für den Bewegungsbeobachtungstest eine lokale stochastische Unabhängigkeit sowie die Homogenität der Items angenommen werden kann. Das **Testverfahren** T_1 berechnet, dass alle 36 Itemkombinationen über der korrigierten Signifikanzgrenze von $\alpha^* = 0,1/36 = 0,003$ liegen und somit alle Items lokal stochastisch unabhängig sind. Mittels der **Teststatistik** T_{1m} wurde ebenfalls kein Itempaar unter dem α^* -Signifikanzniveau identifiziert, so dass Eindimensionalität der neun Items angenommen werden kann. Ebenso kann über das **Testverfahren** T_{1l} gezeigt werden, dass es keine Itemkombinationen gibt, die ein erhöhtes Auftreten des Antwortmusters {11} aufweisen. Eine ausreichende Trennschärfe kann über die **Teststatistik** T_{pbis} nachgewiesen werden, denn die p-Werte aller Items liegen deutlich über der Signifikanzgrenze von $\alpha^* = 0,1/9 = 0,01$, was darauf hindeutet, dass diese Items zu der restlichen Skala der anderen Items passen. Der Test zur Bewegungsbeobachtungsfähigkeit kann somit als ein Instrument bestätigt werden, welches die hohen Anforderungen des Rasch-Modells erfüllt.

5.2.2.3.2 Item- und Personenparameter des Bewegungsbeobachtungstests

Die Bestimmung der Itemschwierigkeit (siehe Tab. 40) zeigt, dass die ersten fünf Items einen negativen Wert haben und somit einer Mehrheit der Testpersonen leichter fallen als die letzten vier Items. Das mit Abstand schwerste Item ist B8 mit einer Itemschwierigkeit von $\beta = 1,006$.

Tab. 40: *Itemschwierigkeit für die 9 Items des Bewegungsbeobachtungstests*

	beta-Wert	SD	lower 0.95 CI	upper 0.95 CI
B1	-0,519	0,277	-1,062	0,025
B2	-0.043	0.257	-0.546	0.461
B3	-0.798	0.295	-1.376	-0.219
B4	-0.433	0.273	-0.967	0.101
B5	-0.700	0.288	-1.265	-0.135
B7	0.381	0.250	-0.109	0.871
B8	1.006	0.257	0.502	1.510
B9	0.793	0.252	0.299	1.287
B10	0.312	0.250	-0.179	0.803

Der durchschnittliche Personenparameter liegt bei einem Wert von $\theta = 0,532$ ($SD = 0,830$) und lässt vermuten, dass die Aufgaben für die Probandengruppe durchaus fordernd ist. Lediglich eine Testperson erreichte den Höchstwert von $\theta = 3,166$ (mit 9 von 9 richtigen Antworten). 15 weitere Probanden erreichten einen Personenparameterwert über Eins, wofür mindestens sieben Items richtig gelöst werden müssen. Den Negativwert von $\theta = -1,358$ (mit 2 von 9 erreichten Antworten) erzielten lediglich zwei Testpersonen, für zwölf weitere Probanden wurde ein negativer Personenwert berechnet (höchs-

tens 4 von 9 Items richtig beantwortet). Die Separation Reliability beträgt 0,1782 (bei einer beobachteten Varianz von 0,7009 (Squared Standard Deviation) und einem Mean Square Measurement Error von 0,576).

Anhand des Histogramms der Person-Item Map (siehe Abb. 94, Person Parameter Distribution) lässt sich erkennen, dass die meisten Testpersonen eine mittlere Personeneigenschaft ($0,24 < \theta < 1,35$) zeigen, während die Personengruppen mit niedrigem Wert ($\theta \leq 0,24$) und mit hohem Personenscore ($\theta \geq 1,35$) deutlich kleiner sind. Die Gruppe mit niedrigem Personenwert ist dabei deutlich größer als die Gruppe mit hohem Personenwert. Die Aufteilung der Itemparameterschwierigkeit auf der latenten Dimension zeigt, dass die Items B1 bis B5 vor allem für die Gruppe mit niedrigem Personenwert passend sind, während sie für die Gruppe mit mittlerem Personenwert zu leicht sind. Die Items B7 bis B10 (mit $1 < \beta < 2$) entsprechen der mittleren Personeneigenschaft der großen Personengruppe. Für die Personengruppe mit hoher Personeneigenschaft (insgesamt sind es 16 Testpersonen mit $\theta > 1$) fehlen passende Items.

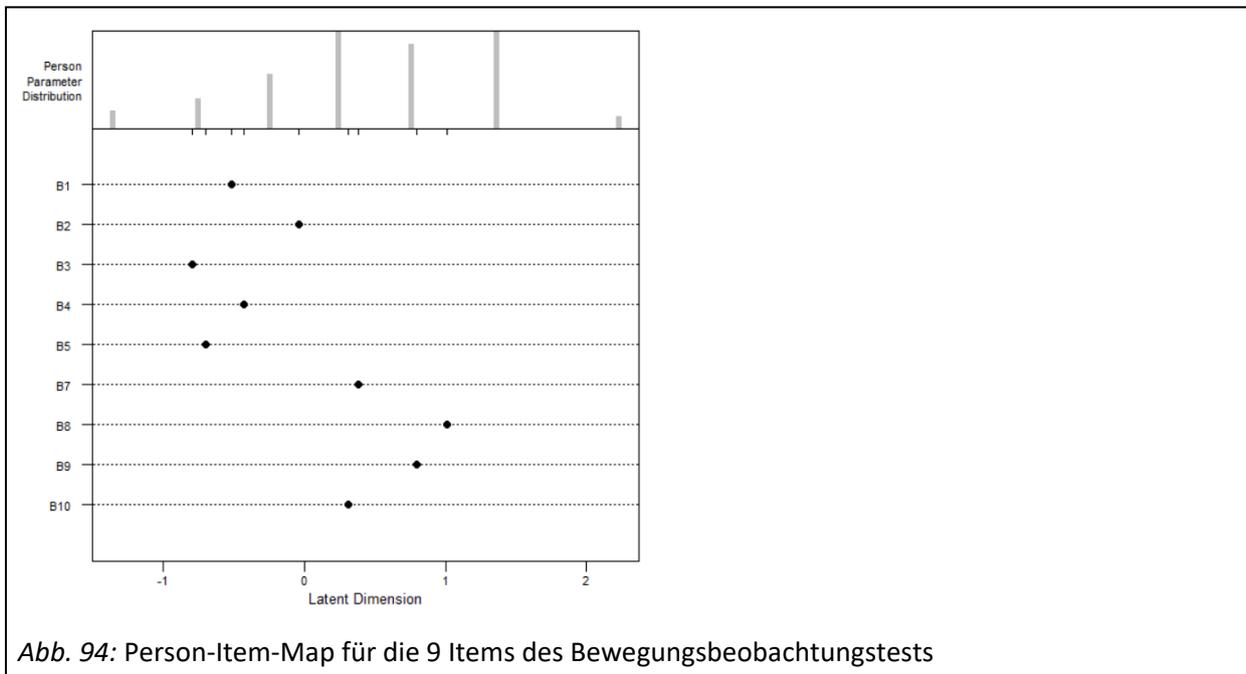


Abb. 94: Person-Item-Map für die 9 Items des Bewegungsbeobachtungstests

5.2.2.3.3 Zusammenhang des TKBV mit dem Bewegungsbeobachtungstest

Berechnet man den Korrelationskoeffizienten nach Spearman zwischen den Personenparametern des TKBV und des Bewegungsbeobachtungstests, so zeigt sich, dass eine gewisse lineare Abhängigkeit zwischen den beiden Testergebnissen besteht ($r = 0,406$). Die Korrelation ist auf dem 0,01-Niveau signifikant. Eine Testperson, die den TKBV erfolgreich bewältigt, schneidet auch beim Bewegungsbeobachtungstest tendenziell besser ab als eine Testperson, die beim TKBV Schwierigkeiten hat. Interessant ist auch das Ergebnis, dass der Korrelationskoeffizient nahezu identisch ist, wenn man die Personenparameter beider Einzelteile des TKBV gesondert berechnet und mit dem Bewegungsbeobachtungstest vergleicht ($r = 0,343$ mit Teil 1, $r = 0,340$ mit Teil 2). Die beiden Einzelteile des TKBV korrelieren aber mit $r = 0,211$ deutlich schwächer miteinander. Dies deutet darauf hin, dass für die erfolgreiche Bewältigung des Bewegungsbeobachtungstests durchaus Fähigkeiten angewendet werden, die für beide Teile des TKBV auch notwendig sind.

5.2.3 Messverfahren der Eigenrealisation

Die Messung der Qualität der Eigenrealisation mittels qualitativer und quantitativer Videoanalyse hat zum Ziel, nach der Datenauswertung für jede Testperson einen Personenparameter vorliegen zu haben, der möglichst objektiv und vergleichbar ist. Die Auswertung der Datenmatrix mit den 39 Items, die mit den Zahlenwerten 0 (Qualitätsstufe 3), 1 (Qualitätsstufe 2) oder 2 (Qualitätsstufe 1) gefüllt sind, erfolgt mit dem ordinalen Rasch-Modell. Der Likelihood-Ratio-Test mit dem Splitkriterium Median liefert ein Ergebnis von $p = 0,427$ und spricht für die Gültigkeit der Annahme des Partial-Credit-Modells. Die geschätzten Itemschwierigkeiten (in diesem Fall basierend auf den vergebenen Qualitätsstufen der 39 Bewegungskriterien) unterscheiden sich demnach zwischen den beiden durch den Median geteilten Gruppen nicht wesentlich und die ermittelten Personenparameter können für den Vergleich der Eigenrealisationsqualität der Testpersonen verwendet werden. Jedoch ist die Wahrscheinlichkeit groß, dass es einige Items gibt, die voneinander abhängig sind, denn einige Bewegungsmerkmale werden zu unterschiedlichen Zeitpunkten im Bewegungsverlauf beurteilt. Die gute oder weniger gute Ausführung des Bewegungskriteriums zu einem vorigen Zeitpunkt wirkt sich dann mit großer Wahrscheinlichkeit auch auf die Beurteilung desselben Merkmals zu einem späteren Zeitpunkt aus. Diese Abhängigkeit kann aber bei einer Bewegungsbeurteilung nicht vermieden werden, bei der viele Bewegungsdetails zu einem Gesamtpersonenparameter zusammengefasst werden sollen.

Wenn man gleich zu Beginn das Gesamtergebnis betrachtet, so liegt der durchschnittliche Personenparameter bei $\theta = 0,390$ ($SD = 0,376$). Die beste Testperson erreicht einen Wert von $\theta = 1,252$, der geringste erreichte Personenparameter liegt bei $\theta = -0,753$. Die Person-Item-Map (siehe Abb. 95) zeigt, dass es von den 39 Beobachtungskriterien nur fünf gibt, bei denen Personen mit schlechterer Bewegungsausführung eine höhere Wahrscheinlichkeit aufweisen, ein besseres Ergebnis zu erzielen als Testpersonen mit einer gesamtdurchschnittlich besseren Eigenrealisation. Es handelt sich hierbei um die Item er13 (Druckfußaufsatz), er18 (Gegenarm), er22 (Stemmfußaufsatz), er27 (Bodenkontakt des Druckbeins) und er32 (Schleifkontakt des Druckbeins).

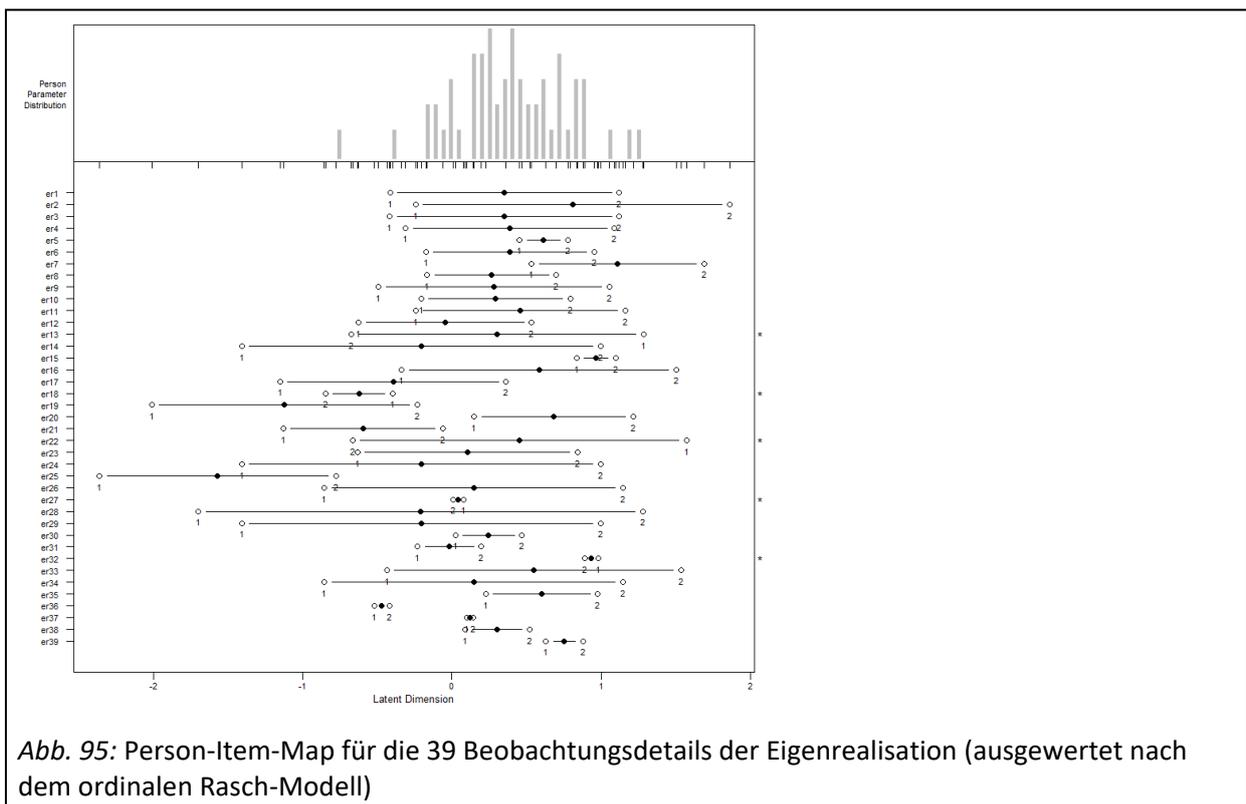


Abb. 95: Person-Item-Map für die 39 Beobachtungsdetails der Eigenrealisation (ausgewertet nach dem ordinalen Rasch-Modell)

Das Histogramm in Abb. 95 zeigt eine kleine Lücke nach dem Personenwerten $\theta = 0,887$ (52 Rohpunkte), wobei nur drei Personen einen Wert von $\theta = 1,062$ (55 Rohpunkte) oder größer erreichen. Dementsprechend fallen drei Beurteilungskriterien auf, bei denen der Durchschnittswellenwert zwischen Qualitätsstufe 1 und 2 (schwarzer Punkt) über $\beta = 0,9$ liegt, was darauf hindeutet, dass diese Bewegungsdetails nur sehr selten technisch richtig ausgeführt wurden. Dazu zählt die aktive Knieführung des Schwungbeins beim Impulsschritt (er7), das Überholen des Druckbeins durch das Stemmbein beim Druckbeinaufsatz (er15) und der Schleifkontakt des Druckbeins im Abwurf (er32). Als besonders gut ausgeführte Bewegungsdetails, deren Schwellenwert von Qualitätsstufe 1 unter $\beta = -0,163$ liegt (nur zwei Personen haben einen entsprechenden Personenparameter unter $\theta = -0,163$) und die somit von vielen Testpersonen technisch sehr gut ausgeführt werden, lassen sich vier Items identifizieren: bei er18 (langer Schwungarm mit Seitrückriss während des Stemmschritts), er19 (Schulterachse in Wurfrichtung in der Wurfauslage), er25 (Brust/Becken frontal in Wurfrichtung in der Bogenspannung) und er36 (Stemmbein gestreckt beim Abfangen) erreichen fast alle Testpersonen mit einer mindestens fünfzigprozentigen Wahrscheinlichkeit Qualitätsstufe 1. Die restlichen Beobachtungskriterien zeigen keine besonderen Auffälligkeiten hinsichtlich ihrer Beurteilung und werden von den Testpersonen entsprechend ihres Leistungsvermögens (ausgedrückt durch den Personenparameter) ausgeführt, ohne dass eine Unter- oder Überforderung zu beobachten wäre, was auch an den durchschnittlich erreichten Punktwerten der einzelnen Beurteilungskriterien gesehen werden kann (siehe Abb. 96).



Abb. 96: Durchschnittlich erreichte Punktzahl der einzelnen Beurteilungskriterien (er1 bis er39)

Interessant erscheint auch der Blick auf die biomechanischen Daten, die in die Bewertung der Eigenrealisation miteinfließen. Die folgenden Messwerte dienen als Beispiel für die Anwendung der sozialen Bezugsnorm zur Festlegung der Intervallbereiche der drei Qualitätsstufen. Für die ermittelte Anlaufgeschwindigkeit zum Zeitpunkt t_1 beispielsweise ergab sich eine Bewertung mit dem Zahlenwert 1 (Qualitätsstufe 2), wenn der biomechanische Messwert #3 zwischen 2,6–3,4 m/s lag. Der Intervallbereich wurde rechnerisch ermittelt, indem jeweils 25 Prozent der Testpersonen über und unter dem Intervallbereich für Qualitätsstufe 2 liegen sollten. Im Fall der Anlaufgeschwindigkeit bei t_1 waren dies jeweils 15 Testpersonen (siehe Abb. 97).

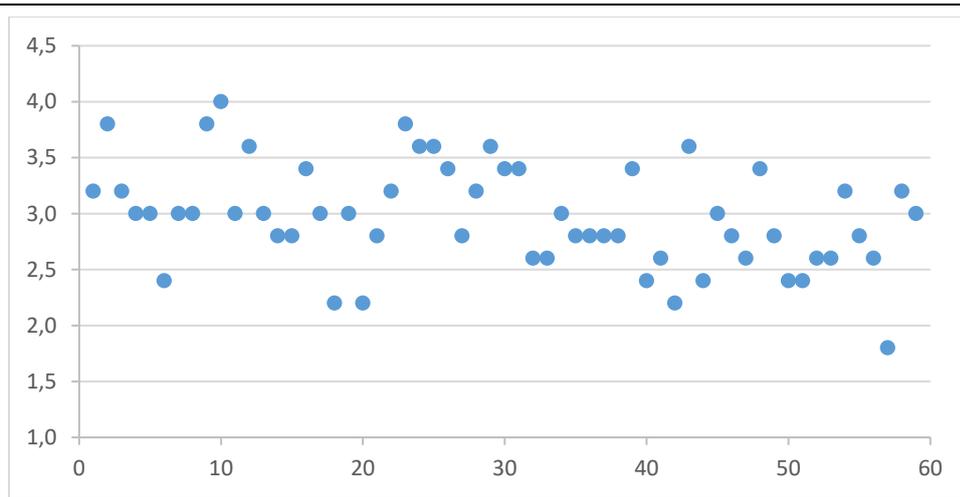
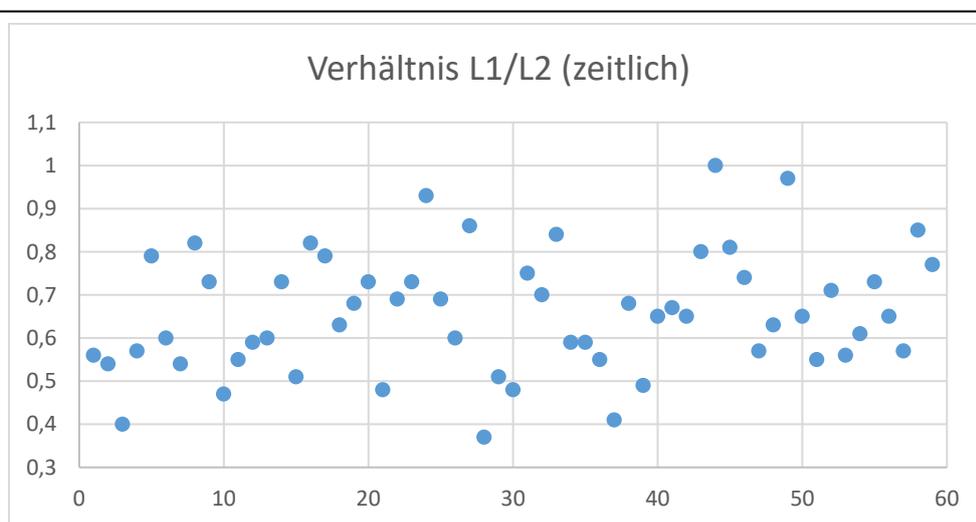


Abb. 97: Biomechanische Messwerte der Anlaufgeschwindigkeit zum Zeitpunkt t_0 (in m/s) der 59 Testpersonen

Als zweites Beispiel dient die Beurteilung des Anlafrhythmus, welcher mittels der biomechanischen Messwerte #4 und #6 in der räumlichen sowie mittels biomechanischer Messwerte #5 und #7 in der zeitlichen Gestaltung von Impuls- und Stemschritt seinen Ausdruck findet. Über den Quotienten aus der Schrittlängengestaltung $L1/L2$ kann beurteilt werden, ob es der Testperson gelingt, den Impulsschritt länger als den Stemschritt auszuführen ($L1/L2 > 1$), während der Quotient aus der Dauer der beiden letzten Anlaufschritte (vom jeweiligen Beginn des Bodenkontakts gemessen) eine zeitliche Einschätzung möglich macht. Die angewendete 25-Prozent-Skalierung ergibt für die Qualitätsstufe 2 folglich einen Intervallbereich für die zeitliche Schrittgestaltung von 0,56-0,74, in denen 50 Prozent der Testpersonen liegen, für die räumliche Schrittgestaltung einen Intervallbereich von 0,88-1,22. Je kleiner der Quotient ist, desto besser ist die technische Anforderung gelungen, den Impulsschritt räumlich lang und den Stemschritt kurz zu setzen bzw. die Flugphase des Impulsschrittes durch die aktive Beschleunigung zu verlängern und zugleich das Setzen der letzten beiden Bodenkontakte möglichst kurz zu halten. Daher deutet ein kleiner Quotientenwert in beiderlei Hinsicht auf eine sehr gelungene rhythmische Ausführung hin, während ein großer Quotientenwert ein aussagekräftiges Zeichen für einen misslungenen Bewegungsrhythmus darstellt (siehe Abb. 98).



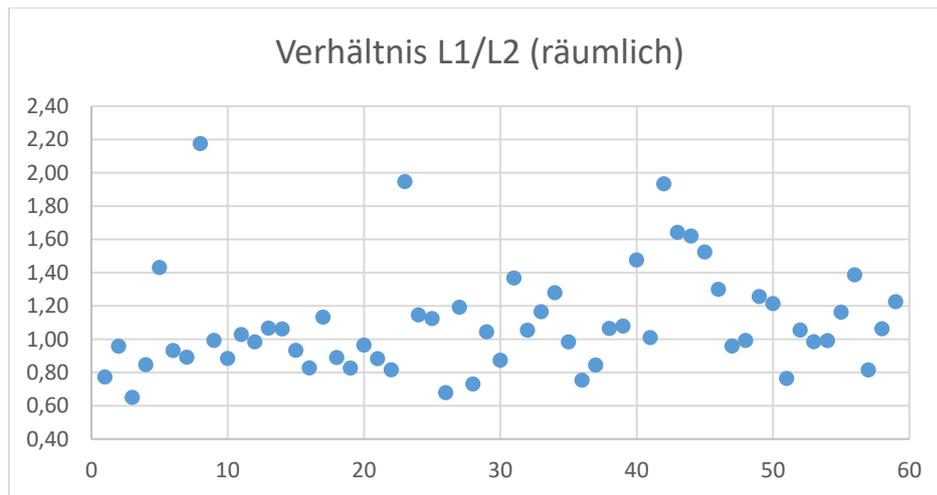


Abb. 98: Zeitliches (oben) und räumliches Verhältnis (unten) der letzten beiden Anlaufschritte

Zur Einschätzung der Anlaufgestaltung wird auch kriteriumsbezogen bewertet, ob es zum Zeitpunkt t_1 im Vergleich zum Zeitpunkt t_3 zu einer Zunahme der Anlaufgeschwindigkeit (Qualitätsstufe 1), zu keiner Veränderung (Qualitätsstufe 2) oder zu einer Abnahme der Anlaufgeschwindigkeit (Qualitätsstufe 3) kommt. Aufgrund leicht auftretender Messungenauigkeit wird hierbei für die Qualitätsstufe 2 ein Intervall von $\pm 0,2$ m/s als gleichbleibende Anlaufgeschwindigkeit akzeptiert. Kleinere negative Differenzen ($v_{t1}-v_{t3} < -0,2$) entsprechen folglich Qualitätsstufe 3, größere positive Differenzen ($v_{t1}-v_{t3} > 0,2$) folglich Qualitätsstufe 1 (siehe Abb. 99).

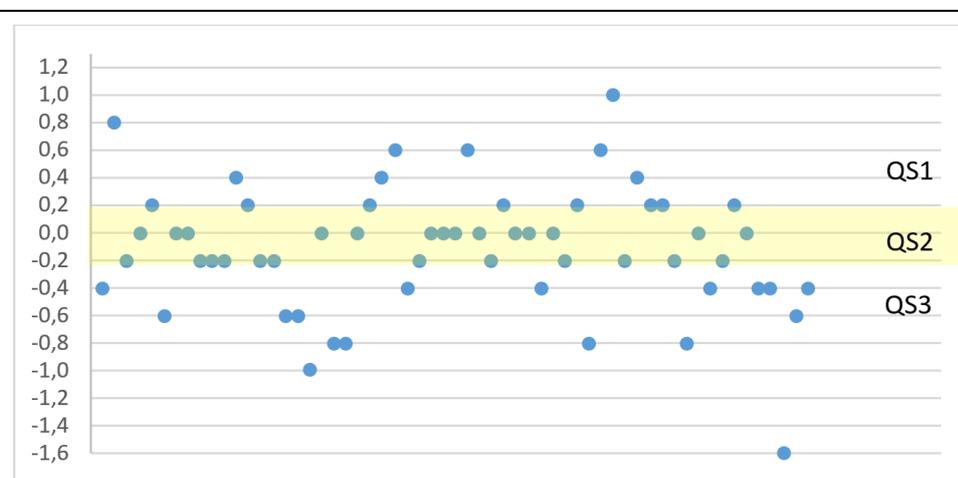


Abb. 99: Differenz der Anlaufgeschwindigkeit zwischen t_1 zu t_3

Eine ebensolche kriteriumsbezogene Auswertung wird für die Beurteilung des Abflugwinkels gewählt. Der Intervallbereich für Qualitätsstufe 1 ist zwar gestützt auf Empfehlungen der leichtathletischen Standardliteratur für den optimalen Abflugwinkel beim Ballwurf, die Intervallbereiche von Qualitätsstufe 2 und 3 sind jedoch vom Verfasser dieser Studie als Kriterien festgelegt. So kommt es, dass 25 Testpersonen das Kriterium für Qualitätsstufe 1 erfüllen, während Qualitätsstufe 3 nur von 14 und Qualitätsstufe 2 von 20 Probandinnen und Probanden erreicht wird (siehe Abb. 100). Die Diagramme aller weiteren biomechanischen Messwerte können im Appendix nachgeschlagen werden.

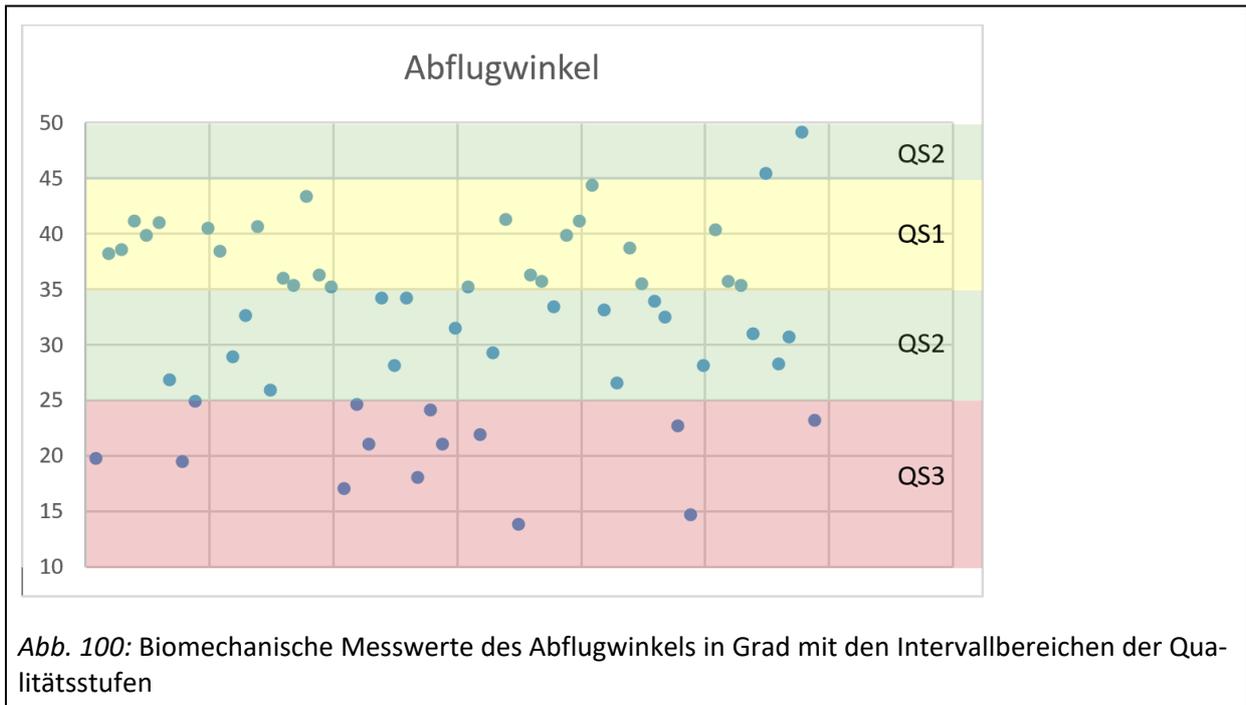


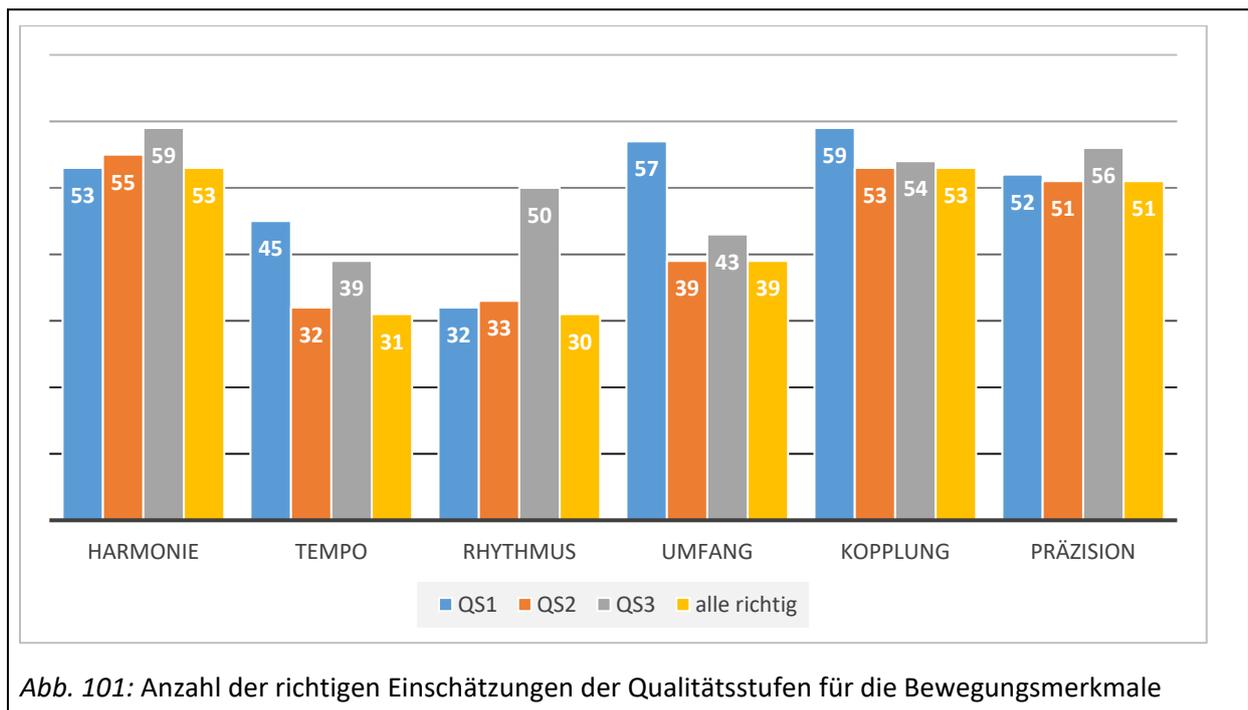
Abb. 100: Biomechanische Messwerte des Abflugwinkels in Grad mit den Intervallbereichen der Qualitätsstufen

5.2.4 Test zur Erfassung der Beurteilungsleistung der Schlagwurfbewegung

Der Beurteilungstest war der letzte Test der zu absolvierenden Messverfahren und wurde im Zeitraum Juni/Juli 2017 durchgeführt. Pro Testtermin konnten bis zu fünf Personen gleichzeitig im E-Learning-Raum U10 des Zentrums für Medien und Kommunikation der Universität Passau den Beurteilungstest absolvieren. Die Testperson wurde zunächst mittels einer Einweisung mit dem Beurteilungsraster (siehe Kapitel 4.5.2) vertraut gemacht, um auf diese Weise die einzelnen Kategorien möglichst gut zu verinnerlichen und die Beurteilung der geforderten Bewegungsmerkmale ohne allzu langes Reflektieren oder Nachdenken durchzuführen. Das Beurteilungsraster mit den Verbalqualifikationen durfte beim Beurteilungstest weiterverwendet werden. Jede der fünf Beurteilungsaufgaben (siehe Kapitel 4.5.3) wurde vor seiner Durchführung mithilfe eines vorgegebenen Textes erklärt und anschließend Raum für Fragen gegeben, um vor der Bearbeitung der ersten Testaufgabe bestehende Unsicherheiten zu klären. Vor Beginn der Beurteilungsaufgaben von Teil A wurde den Testpersonen für jedes Bewegungsmerkmal ein Lehrvideo gezeigt, in welchem die unterschiedlichen Qualitätsstufen der Bewegungsausführung nochmals anhand konkreter Bewegungsbeispiele erklärt wurden. Während des Tests saßen die Testpersonen im Abstand von drei Metern einzeln vor einem Smartboard und bekamen die entsprechenden Bewegungsabläufe als Video präsentiert. Ein Countdown markierte den Beginn jeder Videoaufgabe, um den Beobachter nicht durch einen abrupten Bewegungsbeginn zu überraschen. Nach dem Beobachten der Bewegung trug die Versuchsperson das Ergebnis in einen Auswertungsbogen ein, indem die entsprechende Qualitätsstufe (oder bei Teil C das entsprechende Antwortfeld) angekreuzt wurde.

5.2.4.1 Teil A: Fähigkeit zur Unterscheidung der Qualitätsstufen

Bei Teil A des Beurteilungstests wird von den Testpersonen gefordert, die drei gezeigten Bewegungsausführungen je einer Qualitätsstufe zuzuordnen. Diese Zuordnungsaufgabe ist als Einstieg in die Bewegungsbeurteilung konzipiert, in welcher die Testpersonen die zuvor in einem kurzen Lehrvideo erklärten Qualitätsstufen auf neue Bewegungsbeispiele anwenden müssen. Abb. 101 zeigt die Anzahl der richtigen Einschätzungen der Qualitätsstufen für jedes Bewegungsmerkmal. Betrachtet man die Anzahl der Probanden, die alle drei Beispiele richtig zugeordnet haben (gelber Balken), so fällt auf, dass die Merkmale Harmonie (53), Kopplung (53) und Präzision (51) deutlich erfolgreicher eingeschätzt werden können als die Merkmale Tempo (31), Rhythmus (30) und Umfang (39). Außerdem gelingt es den Testpersonen unabhängig vom Bewegungsmerkmal, qualitativ schlechtere Bewegungsausführungen (Qualitätsstufe 3, grauer Balken) zuverlässiger zu identifizieren als solche der Qualitätsstufe 2 (oranger Balken). Auch die Zuordnungen qualitativ hochwertiger Bewegungsausführungen (Qualitätsstufe 1, blauer Balken) gelingt bei den Merkmalen Harmonie, Rhythmus, Umfang und Präzision nahezu genauso gut und bei den Merkmalen Tempo und Umfang deutlich besser bei Bewegungsausführungen der Qualitätsstufe 2. Dies deutet darauf hin, dass es den Testpersonen hauptsächlich bei den Merkmalen Tempo, Rhythmus und Umfang weniger gut gelingt als bei den Merkmalen Harmonie, Kopplung und Präzision, die Qualitätsstufen 1 und 3 von der Qualitätsstufe 2 genau abgrenzen zu können.



Wenn man Teil A des Beurteilungstests nach dem **dichotomen Rasch-Modell** auswertet, indem ein Proband eine Aufgabe nur dann richtig beantwortet hat, wenn jeder Teilbewegung die richtige Qualitätsstufe zugeordnet wird, so ergibt sich für die Auswertung eine Datenmatrix mit den sechs Items A1 bis A6. Die Codierung der Items erfolgt nach der Maßgabe, dass der Wert 1 nur vergeben wird, wenn alle drei Zuordnungen der Teilaufgabe richtig sind, und der Wert 0 bedeutet, dass mindestens eine Zuordnung falsch war. Der durchschnittliche Personenparameter liegt dann bei einem Wert von $\theta = 1,411$ (SD = 1,466) mit einem von zwölf Testpersonen erreichten Höchstwert von $\theta = 3,236$ (alle

sechs Items richtig beantwortet) und einem einzeln vorliegenden Tiefstwert von $\theta = -3,203$ (keines der sechs Items richtig beantwortet). Die Separation Reliability beträgt 0,4668 (bei einer beobachteten Varianz von 2,1866 (Squared Standard Deviation) und einem Mean Square Measurement Error von 1,1659).

Anhand des Histogramms der Person-Item Map (siehe Abb. 102, Person Parameter Distribution) lässt sich erkennen, dass viele Personen einen hohen Personeneigenschaftswert $\theta > 2$ erreichen und die Anzahl der Testpersonen mit Abnahme des Personenwerts auch kontinuierlich sinkt. Die Aufteilung der Itemparameterschwierigkeit auf der latenten Dimension zeigt, dass die Items A1 (Bewegungsharmonie), A5 (Bewegungskopplung) und A6 (Bewegungspräzision) für fast alle Testpersonen sehr einfach zu lösen sind und bedacht werden muss, dass es nur wenige Probanden gibt, für welche der Anspruch dieser Aufgaben angemessen erscheint. Die Items A2 (Bewegungstempo), A3 (Bewegungsrythmus) und A4 (Bewegungskopplung) hingegen erscheinen für den mittleren Personenbereich ($0 < \theta < 2$) gut lösbar, jedoch fehlen schwierigere Items für den Personenanteil mit hoher Personeneigenschaft. Dies ist jedoch in Anbetracht der Tatsache, dass Teil A des Beurteilungstests als eine einfache Einstiegsübung für die Einübung der Qualitätsstufen konstruiert wurde, nicht überraschend, sondern sogar gewollt. Die Einschätzung der Bewegungsmerkmale in den drei Qualitätsstufen sollte auf diese Weise geübt und für die späteren, schwierigeren Aufgaben gefestigt werden.

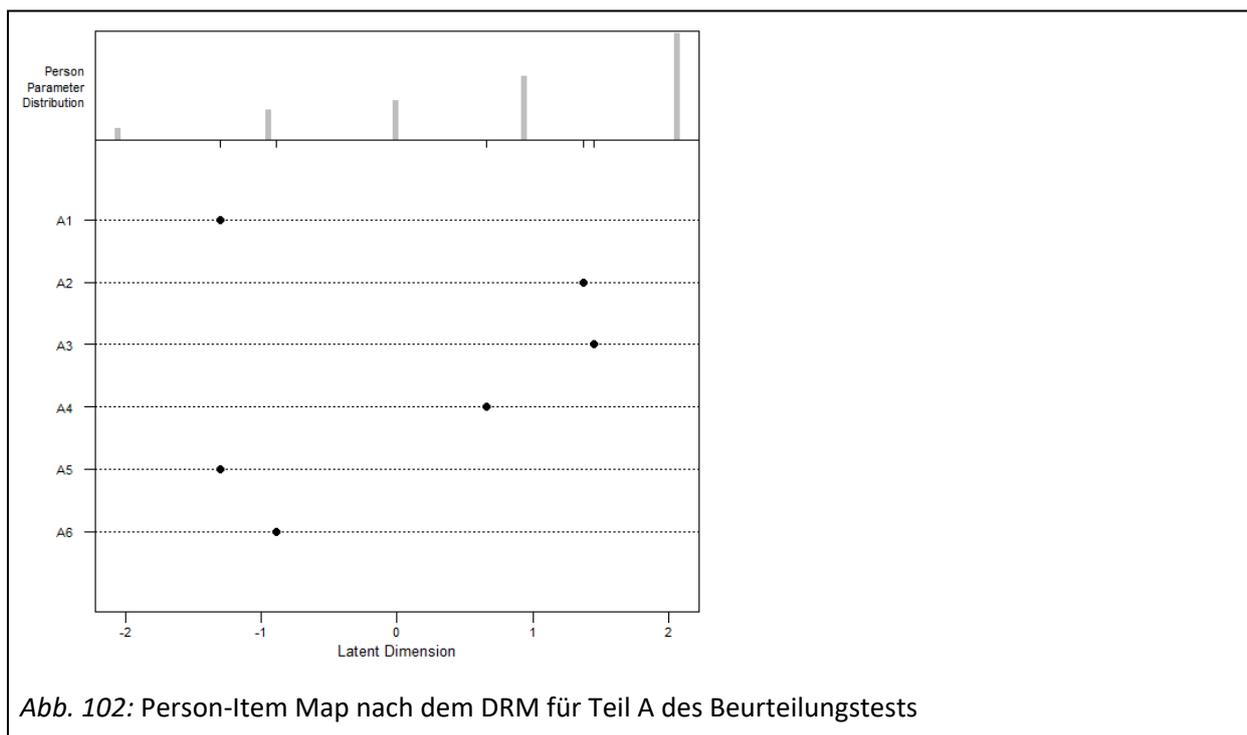


Abb. 102: Person-Item Map nach dem DRM für Teil A des Beurteilungstests

Wenn Teil A nach dem **Partial-Credit-Modell** ausgewertet wird, so kann bei der Codierung ausgedrückt werden, in welchem Maße die Testperson mit den Experten übereinstimmt. Da sich die Experten abgesehen von Item A2 absolut einig waren, wurde für die Items A1, A3, A4, A5 und A6 eine dreifache Übereinstimmung mit dem Wert 2, eine einfache Übereinstimmung mit dem Wert 1 und keine Übereinstimmung mit dem Wert 0 codiert. Da einige Testpersonen die Aufgabe nicht nach Vorgabe des Testleiters bearbeitet haben, dass alle drei Qualitätsstufen pro Item zugeordnet werden müssen, sondern zwei Bewegungen des gleichen Items auch mit der gleichen Qualitätsstufe beurteilt wurden,

konnte der Fall eintreten, dass zwei Bewegungen entsprechend des Expertenurteils beurteilt wurden, die dritte Bewegung jedoch vom Expertenurteil abwich. Dieser Fall wurde auch mit dem Wert 1 codiert. Bei den Items A4, A5 und A6 gibt es keine Testperson, die alle drei Bewegungen abweichend vom Expertenvotum beurteilt hat, weswegen der Wert 0 in der Codierung fehlt. Somit erfolgt die Auswertung auf dichotome Weise.

Da sich für Item A2 (Bewegungstempo) die Experten nicht vollständig einig waren, wurde die Auswertung wie folgt vorgenommen: Als volle Übereinstimmung mit den Experten wurde bewertet, wenn A2.1 mit Qualitätsstufe 2 oder 3, A2.2 mit Qualitätsstufe 1 oder 2 und A2.3 ebenfalls mit Qualitätsstufe 1 oder 3 beurteilt wurde. Wenn A2.1 mit Qualitätsstufe 1 sowie A2.2 oder A2.3 mit Qualitätsstufe 3 beurteilt wurden, so wurde dies als Abweichung vom Expertenurteil betrachtet. Drei Übereinstimmungen wurden mit dem Wert 2 codiert, zwei Übereinstimmungen mit dem Wert 1 und eine oder keine Übereinstimmung mit dem Wert 0. Auf diese Weise ist die Auswertung des Items A2 nach dem ordinalen Rasch-Modell möglich und wird der Tatsache gerecht, trotz unterschiedlicher Experteneinschätzung eine kohärente Codierung vorzunehmen.

Der durchschnittliche Personenwert für Teil A nach dem Partial-Credit-Modell beträgt $\theta = 1,643$ (SD = 1,073), was im Vergleich zum dichotomen Raschmodell ($\theta = 1,411$) eine geringfügige Erhöhung darstellt. 15 Probanden erreichten den Testhöchstwert von $\theta = 2,993$ (9 von 9 erreichbaren Rohpunktwerten). Ein negativer Personenwert (bei 3 bzw. 4 Rohpunkten) lag lediglich bei vier Testpersonen vor. Ob die Auswertung mit dem dichotomen oder dem ordinalen Rasch-Modell vorgenommen wird, spielt bei dieser Aufgabenstellung keine große Rolle. Die Personenfähigkeitswerte beider Modelle korrelieren mit einem Wert von $r = 0,861$. Das Histogramm der Person-Item-Map (siehe Abb. 103) zeigt, dass der größte Teil der Testpersonen mit einem Personenwert $\theta > 1,5$ mit den Testitems unterfordert ist. Auffällig ist außerdem, dass bei Item A1 (Bewegungsharmonie) Testpersonen mit einem geringeren Personenwert eine bessere Beantwortung (d. h. eine stärkere Übereinstimmung mit den Experten) erreichen als Testpersonen mit einem höheren Personenwert.

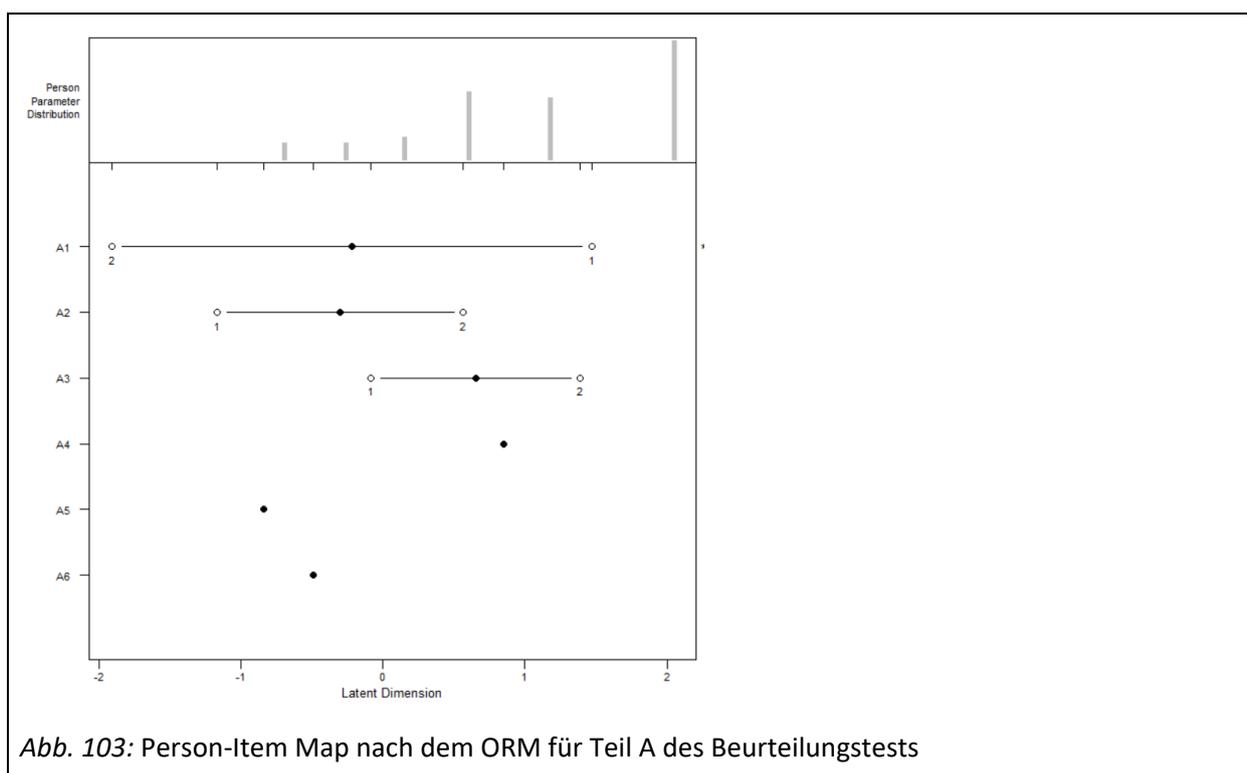


Abb. 103: Person-Item Map nach dem ORM für Teil A des Beurteilungstests

5.2.4.2 Teil B: Fähigkeit zur Einschätzung der Qualitätsstufen

Die Datenauswertung erfolgt in einem ersten Schritt auf Basis des dichotomen Rasch-Modells, wofür jedes einzelne Item entsprechend seiner Übereinstimmung mit dem Hauptexpertenurteil in richtig (Wert 1) oder falsch (Wert 0) codiert wird. Die vergleichende Datenauswertung mit dem Partial-Credit-Modell hat wiederum den Vorteil, dass trotz gewisser Uneinigkeit der Experten alle Items ausgewertet werden können, weil die Antwortmöglichkeiten nicht in richtig oder falsch eingeteilt werden müssen (wie beim dichotomen Rasch-Modell), sondern der Grad der Übereinstimmung mit dem Expertenvotum in die Auswertung miteinbezogen werden kann. Wenn es bei einzelnen Items kein Fehlurteil und auch keine Antwort ‚Keine Einschätzung möglich‘ unter den Testpersonen gibt, so werden diese Items wie dichotome Daten behandelt, wie dies bei den Items B1, B7, B8, B9, B10 und B11 der Fall ist.

Der durchschnittliche Personenparameter **nach dem dichotomen Rasch-Modell** liegt demnach bei einem Wert von $\theta = 1,250$ (SD = 0,681) mit einem von zwei Testpersonen erreichten Höchstwert von $\theta = 2,825$ und einem Tiefstwert von $\theta = -1,166$. Die Separation Reliability beträgt -0,024 (bei einer beobachteten Varianz von 0,6243 (Squared Standard Deviation) und einem Mean Square Measurement Error von 0,6383). Anhand des Histogramms der Person-Item Map (siehe Abb. 104, Person Parameter Distribution) lässt sich erkennen, dass die meisten Testpersonen eine mittlere Personeneigenschaft ($0 < \theta < 1$) zeigen, während die Personen mit niedrigem (d. h. Werten mit $\theta < 0$) und hohem Personenscore (mit $\theta > 1$) vergleichsweise wenige sind. Die Aufteilung der Itemparameterschwierigkeit auf der latenten Dimension zeigt, dass einige Items (mit $\beta < 0$ bei den Items B4, B7, B8, B9 und B10) sehr einfach zu lösen sind und somit die meisten Testpersonen unterfordert sind, weil es nur eine kleine Personengruppe gibt, für welche der Anspruch dieser Aufgaben angemessen erscheint. Dazu gehören vor allem die Bewegungsmerkmale Bewegungsumfang (B7 und B8) und Bewegungskopplung (B9 und B10). Die Items für den mittleren Personenbereich ($0 < \beta < 1$, Items B1, B2, B3 und B12) erscheinen angesichts der Personenverteilung angemessen für einen Großteil der Gruppe, was vor allem auf die Bewegungsharmonie (B1 und B2) zutrifft. Das Item B5 (Bewegungsrhythmus) kann nur von einer kleinen Personengruppe mit einer hohen Personeneigenschaft zuverlässig gelöst werden. Insgesamt scheint die Probandengruppe mit dieser Beurteilungsaufgabe im Durchschnitt gut zurecht zu kommen.

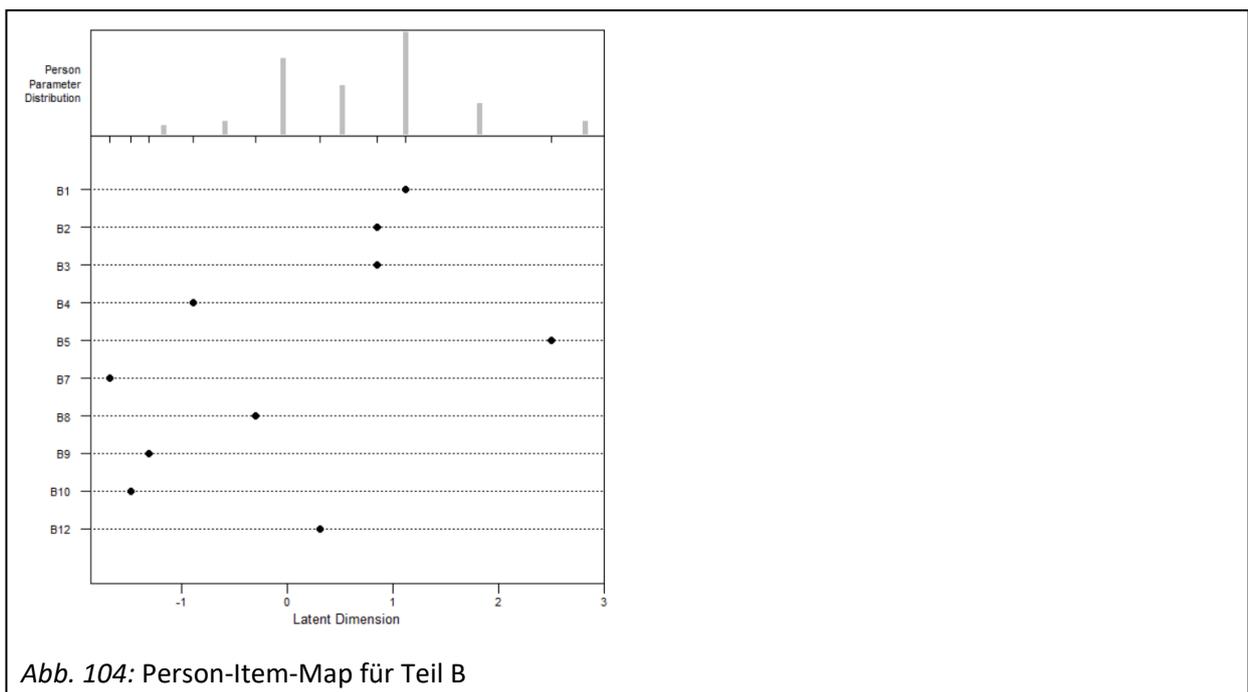


Abb. 104: Person-Item-Map für Teil B

Auffällig an der Auswertung nach dem **Partial-Credit-Modell** von Teil B ist die Tatsache, dass die Hälfte der Items (B1, B7, B8, B9, B10 und B11) derart beantwortet wurden, dass die Beurteilungskategorie 0 (Fehlurteil) keiner der Testpersonen zugewiesen wird (siehe Abb. 105). Somit werden diese Items wie dichotome Daten behandelt und ausgewertet, indem für eine Übereinstimmung mit dem Hauptexpertenurteil der Wert 1 und für eine Übereinstimmung mit dem Nebenexpertenurteil der Wert 0 vergeben wird. Der durchschnittliche Personenwert für Teil B nach dem ordinalen Rasch-Modell beträgt demnach $\theta = 1,400$ ($SD = 0,576$) und stellt im Vergleich zum dichotomen Raschmodell ($\theta = 1,250$) nur eine leichte Erhöhung dar. Ein Proband erreichte den Testhöchstwert von $\theta = 3,270$ (17 von 18 erreichbaren Rohpunkten), drei weitere konnte 16 Rohpunkte erzielen ($\theta = 2,461$). Je eine Testperson erreichte lediglich 8 ($\theta = -0,036$) bzw. 10 Rohpunkte ($\theta = 0,476$). Die Korrelation zwischen den Personenfähigkeitswerten nach dem dichotomen und denen nach dem ordinalen Rasch-Modell ist mit $r = 0,614$ immer noch hoch, aber geringer als bei Teil A. Dies lässt sich damit begründen, dass bei der Auswertung nach dem ordinalen Rasch-Modell keine Items ausgeschlossen wurden.

Das Histogramm der Person-Item-Map für das PCM (siehe Abb. 105) zeigt, dass verhältnismäßig viele Personen mittlere Fähigkeitswerte zwischen $0,983 < \theta < 1,952$ erreichen. Die Items B4 (Bewegungstempo), B5 (Bewegungsrhythmus) und B12 (Bewegungspräzision) zeigen eine überraschende Beantwortung: Im Vergleich zum Nebenexpertenurteil wird bei diesen Items das Hauptexpertenurteil von Personen mit einem geringeren Personenwert mit einer größeren Wahrscheinlichkeit gewählt. Das bedeutet zugleich, dass Personen, die einen hohen Personenwert für Teil B erreichen, bei den Items B4 (Bewegungstempo), B5 (Bewegungsrhythmus) und B12 (Bewegungspräzision) tendenziell nicht mit dem Hauptexpertenurteil übereinstimmen.

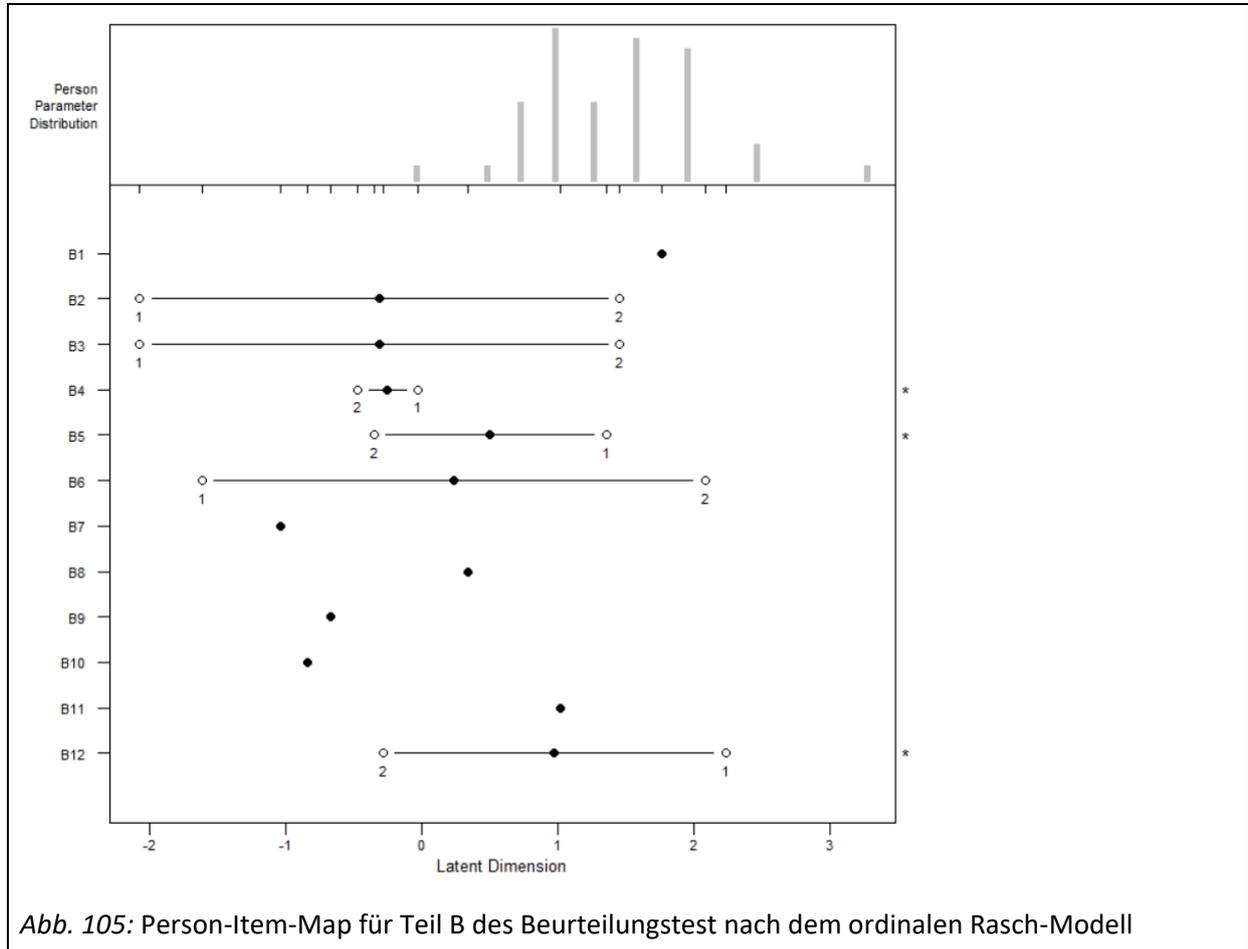


Abb. 105: Person-Item-Map für Teil B des Beurteilungstest nach dem ordinalen Rasch-Modell

5.2.4.3 Teil C: Fähigkeit zum Vergleich von Bewegungsmerkmalen

Bei Teil C besteht die Aufgabe darin, zwei Bewegungsausführungen desselben Sportlers zu beobachten und zu beurteilen, ob sich ein vorgegebenes Bewegungsmerkmal verbessert oder verschlechtert hat oder gleich gut ausgeführt wurde. Wenn bei dichotomer Auswertung die Testperson mit dem Hauptexpertenurteil übereinstimmt, wird die Antwort mit dem Wert 1 codiert, andernfalls mit dem Wert 0. Der durchschnittliche Personenparameter nach dem **dichotomen Rasch-Modell** liegt demnach bei einem Wert von $\theta = -0,150$ (SD = 1,026) und lässt vermuten, dass die Aufgaben für die Probandengruppe fordernd sind. Vier Probanden erreichten den Höchstwert von $\theta = 1,984$ (5 von 6 Items richtig beantwortet), den Negativwert von $\theta = -2,001$ erreichten gleich acht Testpersonen. Die Separation Reliability beträgt 0,069 (bei einer beobachteten Varianz von 1,072 (Squared Standard Deviation) und einem Mean Square Measurement Error von 0,998).

Anhand des Histogramms der Person-Item Map (siehe Abb. 106, Person Parameter Distribution) lässt sich erkennen, dass die meisten Testpersonen eine mittlere Personeneigenschaft ($-1 < \theta < 1$) zeigen, während die Personenanzahl mit niedrigem (d. h. Werten mit $\theta < -1$) und hohem Fähigkeitswert (mit $\theta > 1$) vergleichsweise gering ist. Die Aufteilung der Itemparameterschwierigkeit auf der latenten Dimension zeigt, dass vier Items (mit $-1 < \beta < 1$) passend für den mittleren Personenbereich sind. Item C6 (Bewegungspräzision) ist verhältnismäßig leicht zu lösen, während Item C4 (Bewegungskopplung) eine relativ geringe Lösungswahrscheinlichkeit aufweist. Das Spektrum der auftretenden Personenparameter ist somit gut mit entsprechenden Items abgedeckt.

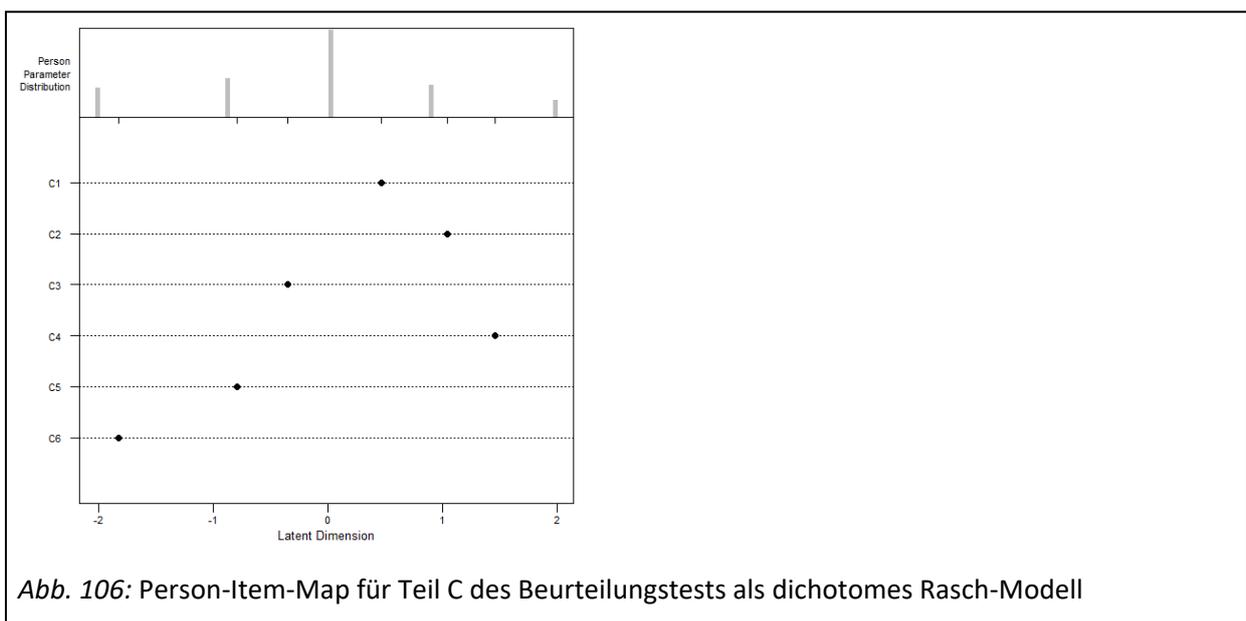


Abb. 106: Person-Item-Map für Teil C des Beurteilungstests als dichotomes Rasch-Modell

Der durchschnittliche Personenparameter nach dem **Partial-Credit-Modell** beträgt für die Items C1 bis C6 $\theta = 0,891$ (SD = 0,689), was im Vergleich zum dichotomen Rasch-Modell ($\theta = -0,150$) eine deutliche Erhöhung darstellt. Zwei Testpersonen erreichten den Maximalpersonenwert von $\theta = 2,717$ (entspricht 11 Rohpunkten), elf weitere erreichten einen Personenwert von $\theta = 1,833$ (entspricht 10 Rohpunkten). Den Tiefstwert von $\theta = -0,624$ erreichten zwei Testpersonen (entspricht 4 Rohpunkten), zwei weitere einen negativen Wert von $\theta = -0,272$ (entspricht 5 Rohpunktwerten). Vergleicht man die Personenwerte nach dem ordinalen Rasch-Modell mit den dichotomen Personenwerten, so ergibt sich ein hoher Korrelationskoeffizient von $r = 0,794$.

Die Person-Item-Map nach dem PCM (siehe Abb. 107) zeigt eine annähernde Normalverteilung der Personenparameter mit einer deutlichen Spitze in der Mitte (20 Testpersonen mit $\theta = 0,811$) und davon abfallenden Häufigkeiten bei niedrigeren und höheren Personenwerten. Die Items C1 (Bewegungsharmonie), C2 (Bewegungstempo) und C4 (Bewegungsumfang) entsprechen der Personenfähigkeit der höchsten Personengruppe. Die Items C3 (Bewegungsrhythmus), C5 (Bewegungskopplung) und C6 (Bewegungspräzision) sind für die Mehrzahl der Testperson unterfordernd. Bewegungskopplung und Bewegungspräzision werden von Personen mit niedrigerem Personenwert besser beantwortet als von Personen mit höherem Personenwert.

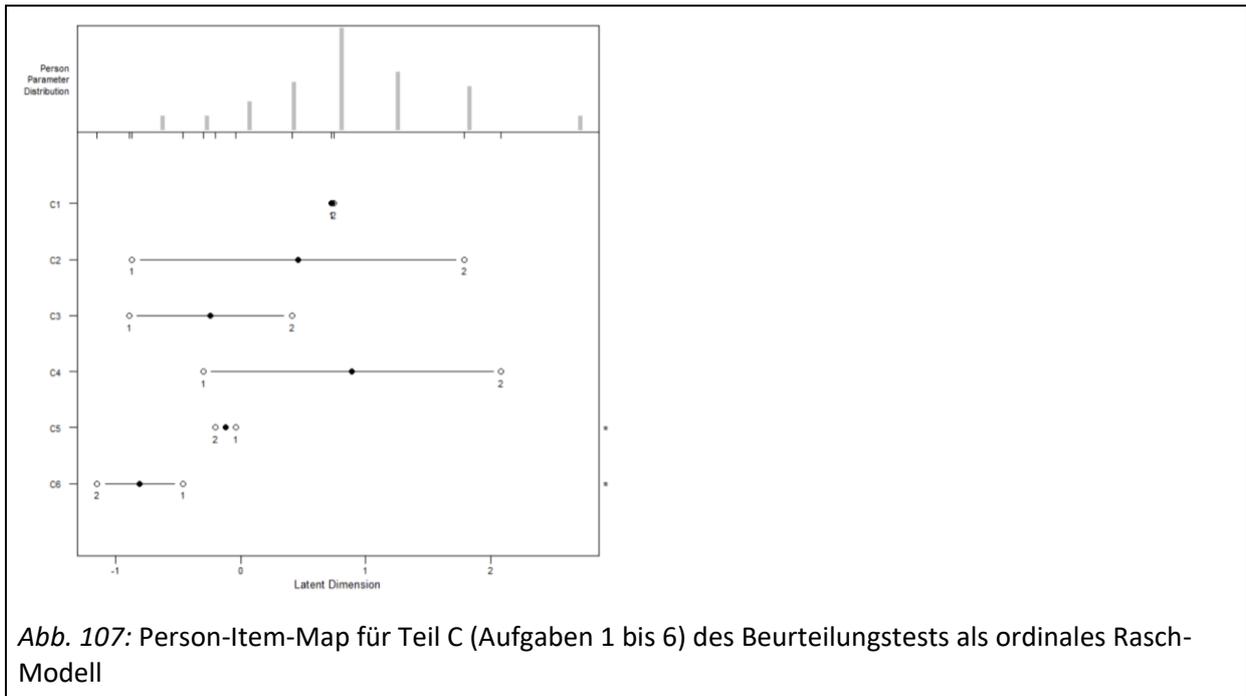


Abb. 107: Person-Item-Map für Teil C (Aufgaben 1 bis 6) des Beurteilungstests als ordinale Rasch-Modell

5.2.4.4 Teil D: Fähigkeit zum Erkennen fehlerhafter Bewegungsmerkmale

Der durchschnittliche Personenparameter für alle zwölf Items nach dem **dichotomen Rasch-Modell** liegt bei einem Wert von $\theta = 0,598$ ($SD = 0,750$) und lässt vermuten, dass vielen Testpersonen die Einschätzung der Bewegungsmerkmale eher schwerfällt. Zwei Probanden erreichten den Testhöchstwert von $\theta = 2,675$ (elf Items richtig eingeschätzt). Der Minimalwert von 4 richtigen Items (und einem Wert von $\theta = -0,795$) wurde bei drei Testpersonen gemessen. Die Separation Reliability beträgt 0,1954 (bei einer beobachteten Varianz von 0,5718 (Squared Standard Deviation) und einem Mean Square Measurement Error von 0,4601). Anhand des Histogramms der Person-Item Map (siehe Abb. 1078, Person Parameter Distribution) lässt sich erkennen, dass die meisten Testpersonen eine mittlere Personeneigenschaft ($0 < \theta < 1$) zeigen, während die Personen mit niedrigem (d. h. Werten mit $\theta < 0$) und hohem Personenscore (mit $\theta > 1$ vergleichsweise wenige sind). Die Gruppe mit hoher Personeneigenschaft ist dabei wiederum größer als die Gruppe mit niedriger Personeneigenschaft. Die Aufteilung der Itemparameterschwierigkeit auf der latenten Dimension zeigt, dass sieben Items (mit $0 < \beta < 1$) passend für den großen mittleren Personenbereich sind. Drei Items fallen in einen leichten Schwierigkeitsbereich ($-1 < \beta < 0$) und können von allen Testpersonen mit hoher Wahrscheinlichkeit gelöst werden. Die Items D1.4 (Bewegungsumfang) und D1.6 (Bewegungspräzision) sind so leicht ($\beta < -1$), dass alle Personen mit der Einschätzung dieser Bewegungsmerkmale unterfordert scheinen. Es fehlen hingegen Items, die für die Gruppe mit hoher Personeneigenschaft passend sind (mit $\beta > 1$).

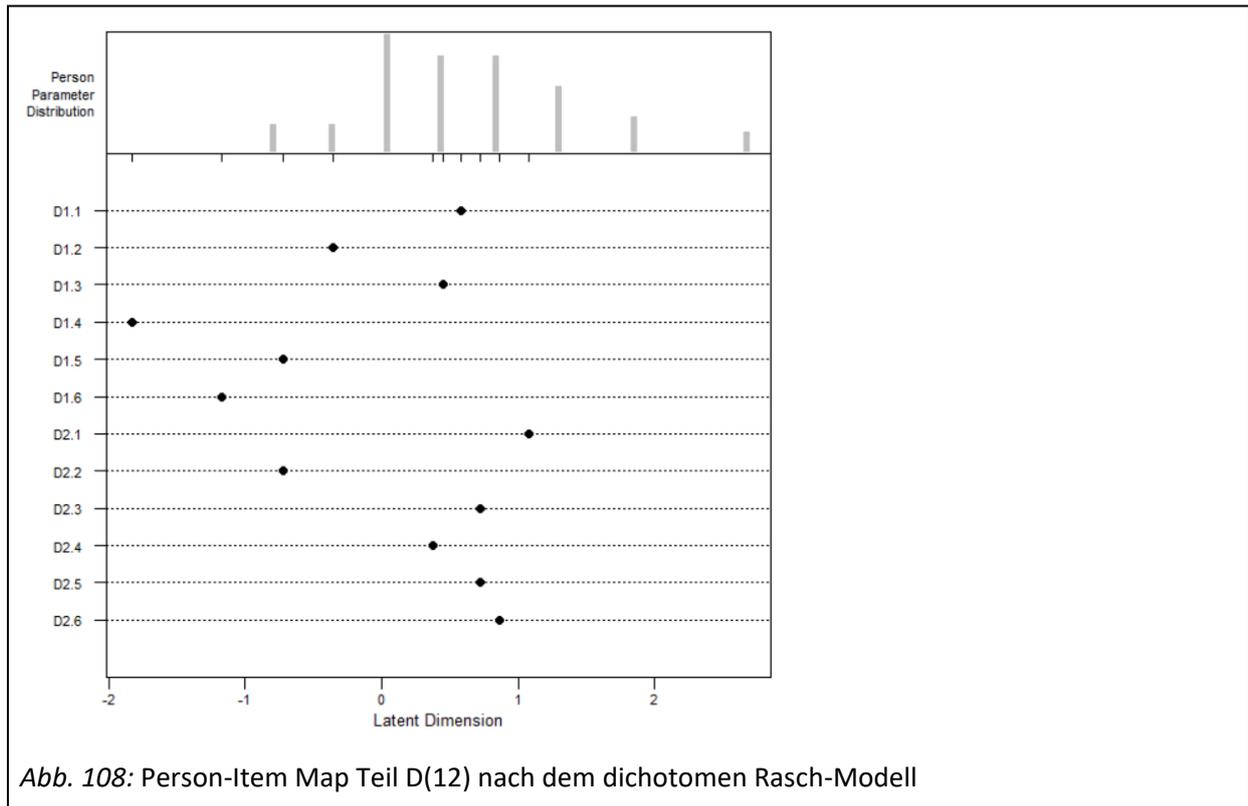


Abb. 108: Person-Item Map Teil D(12) nach dem dichotomen Rasch-Modell

Die Codierung der Daten für Teil D zur Auswertung nach dem **Partial Credit Modell** erfolgt derart, dass nur die Items, welche die Bewegungsharmonie betreffen, nach dem oben beschriebenen Muster gemäß der Übereinstimmung mit den Expertenurteilen codiert werden. Alle anderen Items werden dichotom ausgewertet, weil es bei dieser Aufgabenstellung nur darum geht, Bewegungsmerkmale zu identifizieren, welche der Qualitätsstufe 3 entsprechen. Dass sich die Experten auch nicht einstimmig einig waren, sondern es auch durchaus zwischen den Experten unterschiedliche Meinungen bezüglich des Beurteilungsergebnisses gab, stellt eine gewisse Auswertungsschwäche dieses Aufgabenteils dar, denn letztendlich muss für die Auswertung nach dem Rasch-Modell das Hauptexpertenurteil als ein einstimmiges Urteil vorausgesetzt werden und davon abweichende Experteneinschätzungen bleiben unberücksichtigt.

Wenn man die Teile D1 und D2 des Beurteilungstests nach dem Partial Credit Modell auswertet, dann verbessert sich der durchschnittliche Personenparameter für alle zwölf Items (von $\theta = 0,598$ nach dem dichotomen Raschmodell) auf einen Wert von $\theta = 1,163$ ($SD = 0,720$). Zwei Probanden erreichten den Testhöchstwert von $\theta = 3,214$ (13 der 14 Urteile gemäß dem Expertenurteil), vier weitere konnte zwölf Rohpunkte erzielen ($\theta = 2,396$). Drei Testpersonen erreichten lediglich 6 Rohpunkte ($\theta = -0,078$).

Das Histogramm der Person-Item-Map (siehe Abb. 109) zeigt, dass die Personenfähigkeitswerte der Normalverteilungskurve schon recht ähnlich sind. Die qualitative Einschätzung des Merkmals Bewegungsharmonie (Items D1.1 und D2.1) gelingt Testpersonen mit hoher Personenfähigkeit besser als solchen mit geringerem Personenwert. Insgesamt liegen die Items bei vielen Probanden im machbaren Bereich, einige Testpersonen sind jedoch bei dieser Beurteilungsaufgabe unterfordert.

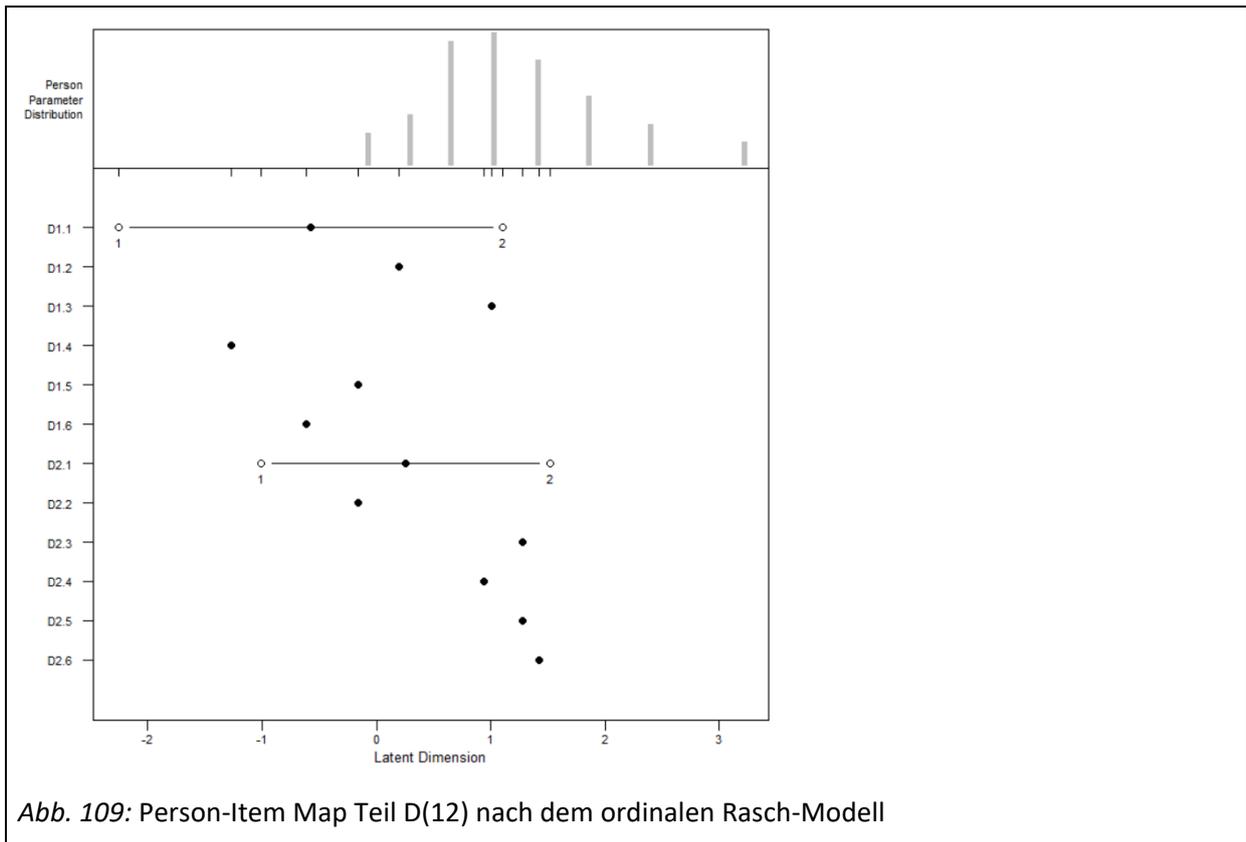


Abb. 109: Person-Item Map Teil D(12) nach dem ordinalen Rasch-Modell

Vergleicht man die Personenwerte der dichotomen Auswertung mit denen der ordinalen Auswertung, so herrscht eine hohe Übereinstimmung ($r = 0,968$). Dies überrascht nicht, weil nur die beiden Items D1.1 und D2.1 ordinal ausgewertet werden können, während die restlichen zehn Items dichotom bleiben.

5.2.4.5 Teil E: Fähigkeit zur synchronen Gesamtanalyse der Bewegungsmerkmale

Die Auswertung von Teil E des Beurteilungstests liefert interessante Ergebnisse. Die Testperson beurteilt alle sechs Bewegungsmerkmale bei nur einem Beobachtungsdurchgang anhand der drei vorgegebenen Qualitätsstufen in Echtzeitpräsentation. Die Bestimmung der Itemschwierigkeit (siehe Abb. 110) zeigt, dass bei Aufgabe E1 die Items E1.1 (Bewegungsharmonie mit $\beta = -0,387$), E1.2 (Bewegungstempo mit $\beta = -0,245$) und E1.4 (Bewegungsumfang mit $\beta = -0,069$) einen negativen Wert aufweisen, während die Items E1.3 (Bewegungsrhythmus mit $\beta = 0,829$), E1.5 (Bewegungskopplung mit $\beta = 0,083$) und E1.6 (Bewegungspräzision mit $\beta = 1,324$) tendenziell schwerer beantwortet werden konnten. Bei Aufgabe E2 weisen die ersten beiden Items (Bewegungsharmonie mit $\beta = 0,453$ und Bewegungstempo mit $\beta = 1,611$) im Gegensatz zu den Merkmalen Bewegungsrhythmus (mit $\beta = -0,647$), Bewegungsumfang (mit $\beta = -0,471$), Bewegungskopplung (mit $\beta = -0,939$) und Bewegungspräzision (mit $\beta = -0,741$) einen positiven Wert auf und sind somit schwerer zu beantworten.

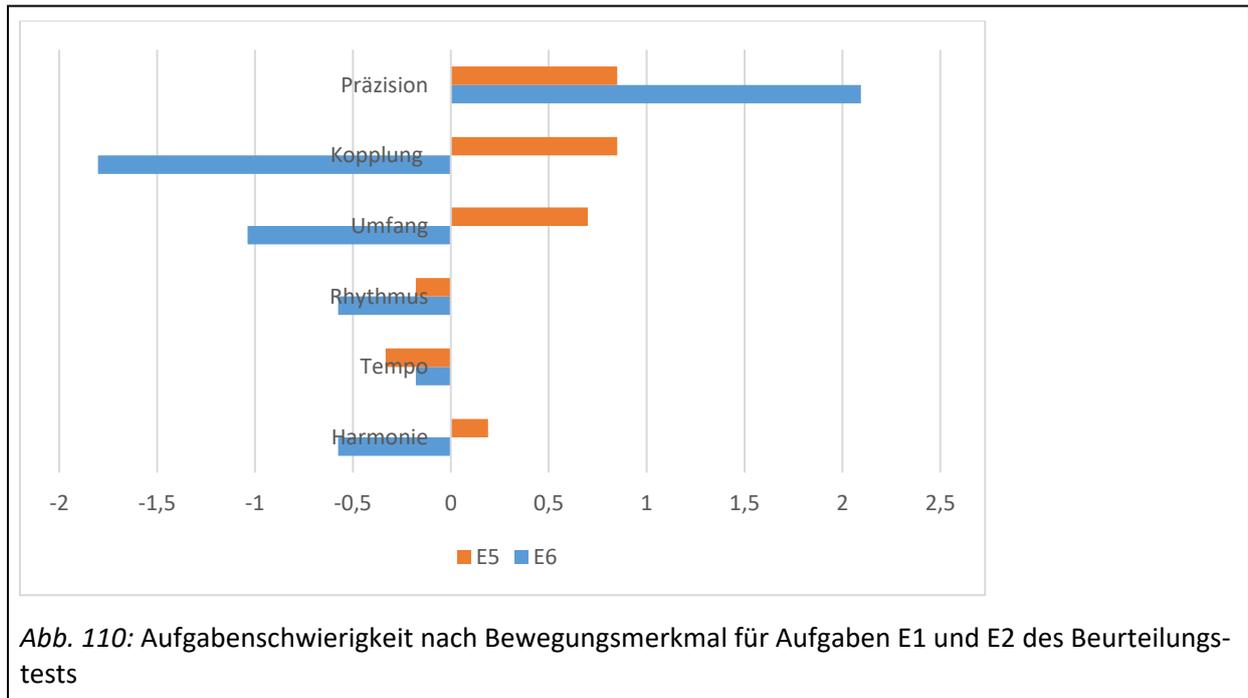


Abb. 110: Aufgabenschwierigkeit nach Bewegungsmerkmal für Aufgaben E1 und E2 des Beurteilungstests

Der durchschnittliche Personenparameter für alle zwölf Items nach dem **dichotomen Rasch-Modell** liegt bei einem Wert von $\theta = 0,367$ ($SD = 1,046$) und lässt vermuten, dass es sich um eine eher schwer zu lösende Aufgabenstellung handelt. Ein Proband erreichte den Testhöchstwert von $\theta = 3,641$ (alle zwölf Items richtig), ein weiterer konnte elf der zwölf Items ($\theta = 2,708$), vier Testpersonen zehn der zwölf Items richtig einschätzen ($\theta = 1,844$). Der Minimalwert von zwei richtigen Items ($\theta = -1,825$) wurde lediglich bei einer Testperson gemessen, eine Testperson konnte kein Item richtig beantworten ($\theta = -2,978$). Anhand des Histogramms der Person-Item Map (siehe Abb. 111, Person Parameter Distribution) lässt sich erkennen, dass die meisten Testpersonen eine mittlere Personeneigenschaft ($-1 < \theta < 1$) zeigen, während die Personen mit niedrigem ($\theta < -1$) und hohem Personenwert ($\theta > 1$) vergleichsweise wenige sind. Die Gruppe mit hoher Personeneigenschaft ist dabei wiederum größer als die Gruppe mit niedriger Personeneigenschaft. Die Aufteilung der Itemparameterschwierigkeit auf der latenten Dimension zeigt, dass fünf Items (mit $\beta < 0,5$) für den großen Teil der Testpersonen mit hoher Wahrscheinlichkeit zu lösen sind. Fünf Items besitzen einen mittleren Schwierigkeitsbereich ($-0,5 < \beta < 1$) und sind für den Großteil der Testpersonen passend. Zwei Items befinden sich in einem hohen Schwierigkeitsbereich ($\beta > 1$) und sind nur von Testpersonen mit hohem Personenwert mit hoher Wahrscheinlichkeit zu lösen.

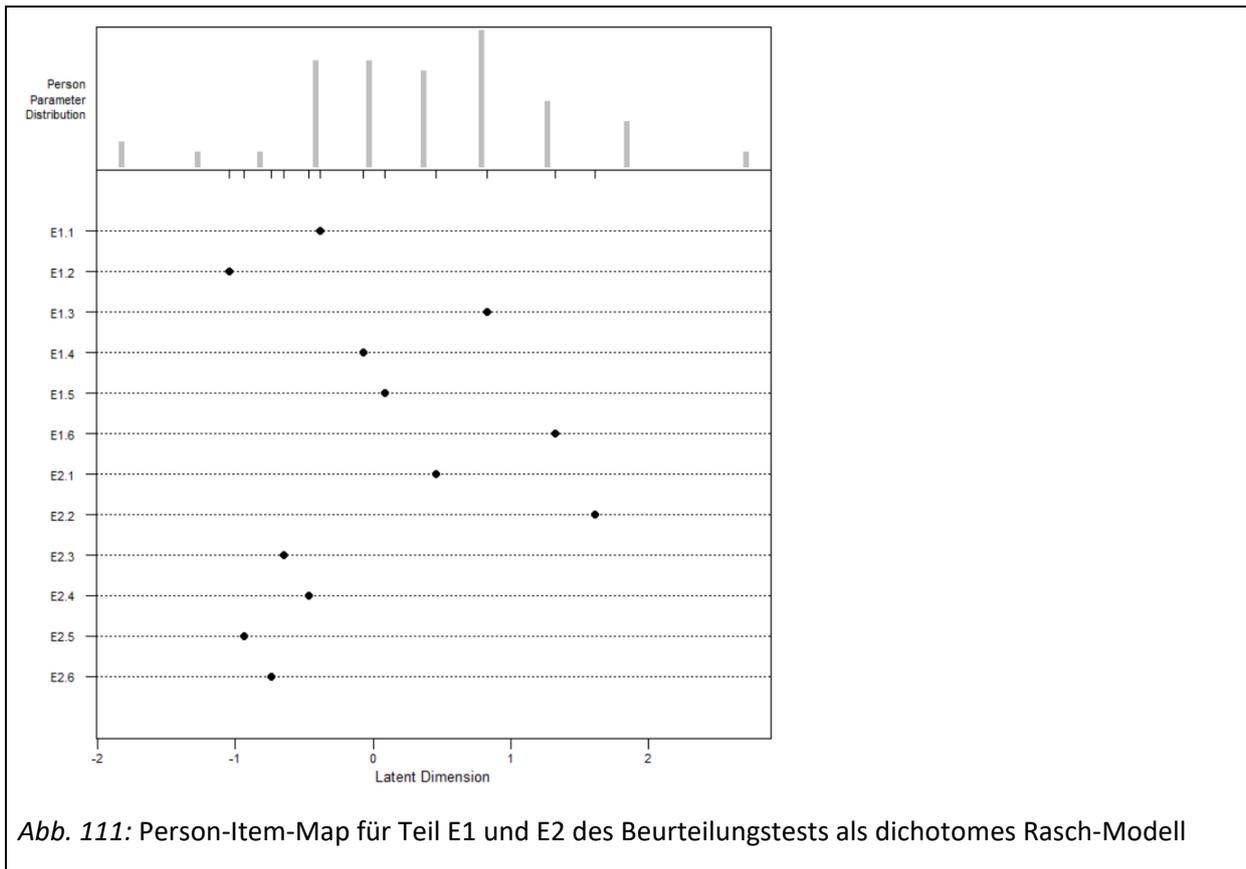


Abb. 111: Person-Item-Map für Teil E1 und E2 des Beurteilungstests als dichotomes Rasch-Modell

Wenn man die Teile E1 und E2 des Beurteilungstests nach dem **Partial-Credit-Modell** auswertet (siehe Abb. 112), dann verbessert sich der durchschnittliche Personenparameter für alle zwölf Items (von $\theta = 0,367$ nach dem dichotomen Raschmodell) auf einen Wert von $\theta = 1,013$ ($SD = 0,803$). Ein Proband erreichte den Testhöchstwert von $\theta = 3,894$ (alle zwölf Items gemäß dem Hauptexpertenurteil mit 24 Rohpunktwerten), ein weiterer konnte elf der zwölf Items gemäß dem Hauptexpertenurteil ($\theta = 3,064$ mit 23 Rohpunktwerten) einschätzen. Eine Testperson erreichte lediglich 8 Rohpunktwerte ($\theta = -0,539$), weitere sechs Testpersonen blieben mit 11 Rohpunktwerten im negativen Wertebereich ($\theta = -0,066$). Die Korrelationskoeffizient zwischen den Personenwerten nach dem ordinalen und nach dem dichotomen Modell liegt bei $r = .747$. Das Histogramm der Person-Item-Map nach dem PCM zeigt, dass die Personenfähigkeitswerte (abgesehen von einem erhöhten Wert bei $\theta = -0,066$) der Normalverteilungskurve schon recht ähnlich sind. Bei einer großen Itemgruppe (E1.2, E1.3, E1.4, E1.5, E2.3, E2.5 und E2.6) stimmen die Personen mit niedrigerem Fähigkeitswert besser mit dem Hauptexpertenurteil überein als Personen mit hohem Fähigkeitswert. Dies kann auf eine hohe Zufallswahrscheinlichkeit bei der Beantwortung dieses Aufgabentyps hindeuten, weil die Testpersonen in der synchronen Einschätzung der vielen Bewegungsmerkmale unsicher oder überfordert sind.

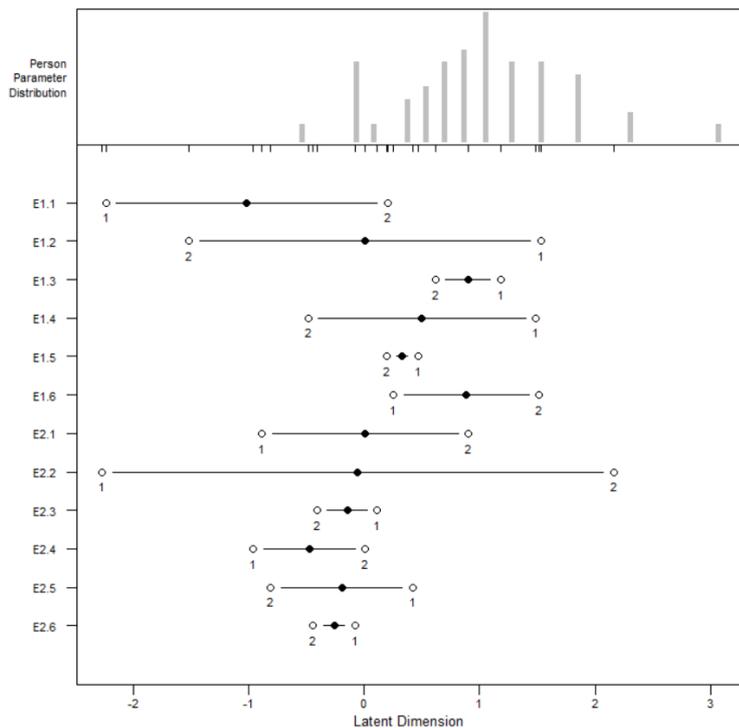


Abb. 112: Person-Item-Map für Teile E1 und E2 des Beurteilungstests als ordinales Rasch-Modell

5.2.4.6 Modellprüfungen

Eine erste Prüfung der grundsätzlichen Modellauswahl erfolgt für die Einzelteile des Beurteilungstests mit dem Likelihood-Quotienten-Test. Bei den Testteilen A, B und E ist sowohl der AIC-Index als auch der BIC-Index für das Modell mit der Annahme der identischen Ladungen (Rasch-Modell) geringer als für das 2PL-Modell mit frei geschätzten Ladungen. Dies spricht sehr begründet dafür, dass das Rasch-Modell die Daten besser erklärt. Für die Teile C und D wird die Bevorzugung des Rasch-Modells zwar nicht durch den AIC-Index, dafür aber durch den aussagekräftigeren BIC-Index bestätigt, weshalb die Datenauswertung des gesamten Beurteilungstests mit dem Rasch-Modell gerechtfertigt erscheint (siehe Tab. 41).

Tab. 41: AIC- und BIC-Index für das Rasch- und das 2PL-Modell für die Einzelteile des Beurteilungstests

		Rasch-Modell	2PL-Modell
Teil A	AIC	357,92	359,79
	BIC	372,46	384,72
Teil B	AIC	628,88	629,97
	BIC	651,75	671,52
Teil C	AIC	418,08	414,03
	BIC	432,62	438,96
Teil D	AIC	871,81	868,97
	BIC	898,82	918,83
Teil E	AIC	881,61	882,30
	BIC	908,62	932,16

In einem zweiten Schritt wird die Gültigkeit des Rasch-Modells mittels der quasi-exakten Testverfahren überprüft. Die **globale Teststatistik T_{10}** zeigt lediglich für Teil D für beide Splitkriterien auf Anhieb einen nicht-signifikanten Wert (siehe Tab. 42) und deutet somit auf die Modellgültigkeit für die entsprechenden Subgruppen hin. Bei Teil A zeigt das Splitkriterium Median mit $p = 1$ an, dass in der oberen Fähigkeitsgruppe (oberhalb des Medians) alle 12 Personen auf jedes Item richtig geantwortet haben und somit alle berechneten Test-Statistiken den gleichen Wert erhalten. Das Teilungskriterium Median ist somit nicht anwendbar. Dafür kann für das Splitkriterium Mittelwert (Mean) ein nicht-signifikanter Wert von $p = 0,717$ berechnet werden, was ebenso wie das Splitkriterium Geschlecht mit $p = 0,45$ die Modellgültigkeit bestätigt. Die Teile B und C zeigen jedoch nicht-signifikante Wert für das Splitkriterium Gender, bei Teil E erweist sich das Kriterium Median als nicht-signifikant. Die Konformität mit dem Rasch-Modell muss somit speziell für die Teile B, C und E auf Itemebene genauer untersucht werden.

Tab. 42: *Signifikanzwerte der Teststatistik T_{10} für die Splitkriterien Median, Gender und Mean für die Einzelteile des Beurteilungstests*

	Median	Gender	Mean
Teil A	1	0,45	0,717
Teil B	0,215	0,011	0,074
Teil C	0,316	0	0,036
Teil D	0,135	0,712	
Teil E	0,039	0,1	

Für Teil A bestätigt die **Teststatistik T_4** auf Itemebene die Gültigkeit des Rasch-Modells, weil sowohl für das Splitkriterium Mittelwert als auch für das Splitkriterium Geschlecht keines der sechs Items bedeutsam vom Rasch-Modell abweicht. Die Bewegungsmerkmale Kopplung (A5) und Präzision (A6) werden von der Personengruppe mit Fähigkeitswert oberhalb des Mittelwerts durchgängig richtig beantwortet. Der hohe p-Wert des Bewegungsmerkmals Umfang (A4 mit $p = 0,968$) für das Splitkriterium Geschlecht bestätigt sich bei der Alternativprüfung nicht ($p = 0,12$), weshalb man nicht davon ausgehen muss, dass bei der Einschätzung des Merkmals Bewegungsumfang das Geschlecht eine Rolle für die Lösungswahrscheinlichkeit spielt. Für Teil B erklärt sich die fehlende globale Modellgültigkeit für das Splitkriterium Geschlecht durch die nicht-signifikanten Items B5 (Bewegungsrhythmus mit $p = 0,007$) und B9 (Bewegungskopplung mit $p = 0,009$), für Teil C analog durch das Item C5 (Bewegungskopplung mit $p = 0,002$). Die zwölf Items von Teil D sind für das Splitkriterium Geschlecht alle modellkonform, während für das Splitkriterium Mean das Item D2.5 (Bewegungskopplung) mit $p = 0,026$ eine Modellverletzung anzeigt. Für Teil E ist das Item E2.1 (Bewegungsharmonie mit $p = 0,004$) für das Splitkriterium Mittelwert sowie das Item E1.2 (Bewegungstempo mit $p = 0,031$) für das Splitkriterium Geschlecht auffällig und spricht für eine unerwünschte Subgruppeninvarianz.

Die **Teststatistik T_{md}** kann für alle Testteile Itemhomogenität nachweisen, wobei für Teil D bei einer Signifikanzgrenze von $\alpha^* = 0,1/2 = 0,05$ (Bonferroni-Korrektur mit zwei Teilungskriterien) die Itemsubgruppen nur knapp als homogen eingestuft werden können. Der nicht-signifikante Wert der globalen **Teststatistik T_{11}** (bei einer Signifikanzgrenze von $\alpha = 0,1$) spricht dafür, dass für die Teile A, B und C

auch eine stochastische Unabhängigkeit vermutet werden kann. Die Teile D ($p = 0,009$) und E ($p = 0,061$) weisen hier eine deutliche Unterschreitung der Signifikanzgrenze auf (siehe Tab. 43) und müssen auf Itemebene genauer bezüglich stochastisch abhängiger Itemkombinationen untersucht werden.

Tab. 43: Ergebnisse der Teststatistiken T_{md} und T_{11}

	T_{md}	T_{11}
Teil A	0,784	0,385
Teil B	0,723	0,466
Teil C	0,298	0,125
Teil D	0,059	0,009
Teil E	0,881	0,061

Das **Testverfahren T_1** konnte für Teil A bei einer Signifikanzgrenze von $\alpha^* = 0,1/15 = 0,0067$ keine signifikanten Abhängigkeiten zwischen den insgesamt 15 Itemkombinationen feststellen und bestätigt somit, dass die Items lokal stochastisch unabhängig sind. Für Teil B kann für das Itempaar B1-B2 mit einem nicht-signifikanten Wert von $p = 0,004$ ($\alpha^* = 0,1/45 = 0,002$) und für Teil C für das Itempaar C2-C4 mit einem nicht-signifikanten Wert von $p = 0,037$ ($\alpha^* = 0,1/36 = 0,03$) nur knapp stochastische Unabhängigkeit festgestellt werden. Für Teil D unterschreitet das Itempaar D5-D6 als einziges die korrigierte Signifikanzgrenze von $\alpha^* = 0,1/66 = 0,0015$ und gilt als abhängig, während für Teil E alle Itemkombinationen als stochastisch unabhängig gelten können. Die **Teststatistik T_{1m}** findet für die Teile A, B, C, D und E bei keinen Itemkombinationen einen Nachweis für Multidimensionalität, so dass Eindimensionalität der jeweiligen Items bestätigt werden kann. Auch ein Lerneffekt durch das sukzessive Beantworten der entsprechenden Items kann mit dem **Testverfahren T_{11}** für alle Teile des Beurteilungstests widerlegt werden, weil es abgesehen vom Itempaar B1-B2 (mit $p = 0,001$ bei $\alpha^* = 0,1/45 = 0,002$) keine signifikanten Itemkombinationen gibt, die ein erhöhtes Auftreten des Antwortmusters {11} nachweisen können. Lediglich die Itemkombinationen C2-C4 (mit $p = 0,034$ bei $\alpha^* = 0,1/36 = 0,03$) und D1.5-D1.6 (mit $p = 0,003$ bei $\alpha^* = 0,1/66 = 0,0015$) erfüllen die Anforderungen nur knapp.

Als letztes kann noch für alle Items der einzelnen Testteile eine ausreichende Trennschärfe mittels der **Teststatistik T_{pbis}** bestätigt werden. Die p-Werte aller Items von Teil A liegen über der Signifikanzgrenze von $\alpha^* = 0,1/6 = 0,016$ (wenngleich das Item A4 mit $p = 0,056$ relativ knapp), was darauf hindeutet, dass alle Items zu der restlichen Skala der anderen Items passen. Ebenso gilt dies für die Items von Teil B für $\alpha^* = 0,1/10 = 0,01$, wobei das Item B12 mit $p = 0,043$ knapp nicht-signifikant ist, Teil C für $\alpha^* = 0,1/6 = 0,0167$, Teil D für $\alpha^* = 0,1/12 = 0,008$ (wenngleich das Item D2.3 mit $p = 0,02$ relativ knapp über der Grenze liegt) und Teil E für $\alpha^* = 0,1/12 = 0,008$ (mit dem knappen Item E1.4 mit $p = 0,017$).

Es lässt sich zusammenfassend sagen, dass alle quasi-exakten Tests die Rasch-Konformität aller Testteile bis auf wenige Ausnahmen bestätigen. Die Fälle, welche eine Modellverletzung anzeigen, können aufgrund ihrer geringen Auswirkung auf die globalen Modellannahmen vernachlässigt werden, und die problematischen Items oder Itemkombinationen können im Beurteilungstest verbleiben.

6 Diskussion

Die Diskussion der Ergebnisse kann in dieser Arbeit nicht einfach anhand der Verifizierung oder Falsifizierung der vor Beginn dieser Studie gebildeten Hypothesen erfolgen, weil sich im Laufe der Forschungsarbeit Entwicklungen und Erkenntnisse abgezeichnet haben, die in dieser Form nicht vorhersehbar waren. Insofern dienten die Hypothesen und vermuteten Zusammenhänge als erster Zugang zur schwierigen Forschungsfrage, mögliche Einflussfaktoren auf die Beurteilungsleistung sportlicher Bewegungen zu identifizieren. Da der Erkenntnisstand in dieser sportwissenschaftlichen Fragestellung äußerst gering ist, können die Hypothesen als wenig theoriegeleitet gelten, sondern sind vorwiegend explorativ entwickelt worden, um neuen Erkenntnisgewinn zu generieren. Dementsprechend besteht die Leistung dieser Studie darin, sich dem Phänomen der Bewegungsbeurteilung empirisch zu nähern, indem ein Testverfahren entwickelt wurde, welches die Beurteilungsleistung am Beispiel der leichtathletischen Schlagwurfbewegung messbar macht. Ebenso wurden weitere Testverfahren konstruiert, um mögliche Einflussfaktoren auf die Beurteilungsleistung untersuchen und in ihren unterschiedlichen Wirkungen identifizieren zu können. Die Interpretation der Ergebnisse muss dabei noch mit einem sehr großen Fokus geführt werden, um Zusammenhänge zwischen den untersuchten Variablen erkennen und richtig darstellen zu können. Dennoch können einige statistisch gut abgesicherte Resultate präsentiert und Schlussfolgerungen gezogen werden, welche wiederum als Ausgangspunkt für neue Forschungsfragen dienen können, um das Phänomen der Bewegungsbeurteilung noch umfassender verstehen zu können.

Zugleich gilt zu erwähnen, dass statistisch nicht signifikante Ergebnisse nicht per se bei der Diskussion unberücksichtigt bleiben müssen. Bühner und Ziegler (2009, S. 666) bestätigen jedenfalls, dass speziell bei kleinen Stichproben nicht signifikante Prädiktoren durchaus beibehalten werden können, wenn sie theoriegeleitet in eine solche Analyse aufgenommen wurden und man sich im Klaren darüber ist, ab welchem B- oder β -Gewicht ein Prädiktor nützlich ist. Die Regressionsgewichte werden nämlich bei kleinen Stichproben nur ungenau geschätzt und es ist schwerer, einen Effekt bei geringer Teststärke in der Grundgesamtheit zu entdecken. Es ist dann zu empfehlen, diese Prädiktoren in einer weiteren Untersuchung erneut zu prüfen, um sicher zu gehen, dass man keine Prädiktoren berücksichtigt, die in der Population eigentlich keinen Effekt haben sollten. Insofern werden auch in dieser Studie gelegentlich nicht signifikante Regressionskoeffizienten bei ausreichender Effektstärke für die inhaltliche Begründung einzelner Ergebnisse herangezogen, jedoch nicht ohne zu erwähnen, dass es sich hierbei um sehr hypothetische Aussagen handelt.

6.1 Messung der Beurteilungsleistung

Zu Beginn der folgenden Diskussion muss festgehalten werden, dass das Konstrukt Beurteilungsleistung schwierig zu operationalisieren ist. Das Beurteilen sportlicher Bewegungen vollzieht sich in der Realität in sovielen verschiedenen Kontexten und Situationen, dass nicht einfach von einer allgemeinen Beurteilungsfähigkeit ausgegangen werden kann, sondern vor einer Untersuchung genau festgelegt werden muss, wie die Beurteilungsaufgabe konkret aussieht. Eine große Leistung dieser Arbeit besteht darin, dass fünf verschiedene Beurteilungsaufgaben konstruiert wurden, die in ihrer Schwierigkeit und Komplexität schrittweise zunehmen. Dies kann anhand der nach dem ordinalen Rasch-Modell ermittelten Personenparameter der 59 Probandinnen und Probanden beobachtet werden, die von

der einfachsten Beurteilungsaufgabe, in welcher die drei Qualitätsstufen einer Bewegungsausführung lediglich unterschieden werden müssen (Teil A mit $\theta_A = 1,643$), bis hin zur synchronen Gesamtanalyse aller sechs Bewegungsmerkmale (Teil E mit $\theta_E = 1,013$) kontinuierlich abnehmen (siehe Kapitel 5.1.1, Abb. 76). Das Beurteilen eines einzelnen Bewegungsmerkmals (Teil B mit $\theta_B = 1,400$) konnten die Testpersonen dabei besser bewältigen als das Identifizieren sehr defizitär ausgeführter Bewegungsmerkmale (Teil D mit $\theta_D = 1,163$). Lediglich die vergleichende Beurteilung eines Bewegungsmerkmals bei zwei Bewegungsausführungen desselbe Sportlers (Teil C mit $\theta_C = 0,891$) entspricht nicht der Hierarchie, welche in der Testkonstruktion angenommen wurde, denn dieser Aufgabentyp stellte die höchsten Anforderungen an das Beurteilungsvermögen der Testpersonen und wurde von diesen am unzuverlässigsten gelöst.

Im Zuge der Datenauswertung bestätigte sich auch die Annahme, dass es keinen Erkenntnisgewinn darstellt, aus den fünf Beurteilungsaufgaben einen gemeinsamen Beurteilungsindex zu bilden, weil die fünf Aufgaben zu unterschiedliche Anforderungen an die beurteilende Testgruppe stellen und die sich ergebenden Korrelationen mit den untersuchten Einflussfaktoren für einen solchen Gesamtindex schwach und nicht-signifikant sind. Durchaus plausibel erscheint es jedoch, die Beurteilungsaufgaben in zwei Gruppen zusammenzufassen: die Teile A und B stellen relativ eindimensionale Beurteilungsaufgaben dar, in welchen lediglich ein Bewegungsmerkmal beobachtet und den Qualitätsstufen zugeordnet werden muss, während die Teile C, D und E wesentlich komplexere Beurteilungsprozesse erfordern, in denen die Qualität zweier Bewegungsausführungen miteinander verglichen oder mehrere Bewegungsmerkmale gleichzeitig wahrgenommen werden müssen. Fasst man die Teile A und B zu einem Index (BT_AB) zusammen, so erklären die beiden signifikanten Prädiktoren biomechanisches Fachwissen und Eigenrealisation immerhin 15,6 Prozent der Beurteilungsleistung der Teile A und B, wobei das biomechanische Fachwissen positiv und die Eigenrealisation negativ auf den entsprechenden Personenparameter wirkt. Der gemeinsame Index der Teile C, D und E (BT_CDE) zeigt ebenfalls zwei signifikante Korrelationen mit der Eigenrealisation und Teil 1 des TKBV, welche 12,1 Prozent der Beurteilungsleistung vorhersagen, wobei sich lediglich die Eigenrealisation als signifikanter Prädiktor erweist (siehe Kapitel 5.1.6.6). Wenn man diese Ergebnisse nun ganz stark vereinfachen will, so lässt sich die Tendenz erkennen, dass für die Lösung einfacher Beurteilungsaufgaben das Fachwissen eine entscheidene Rolle spielt, während eine hohe Qualität der Eigenrealisation der zu bewertenden Bewegung das Lösen komplexer Beurteilungsaufgaben begünstigt.

Interessant könnte es sein, diese Feststellung mit der angewandten Sehstrategie der Beurteilenden (siehe Kapitel 2.2.2) in Verbindung zu bringen. Bei einfachen Beurteilungsaufgaben reicht es aus, eine bestimmte Körperregion foveal zu verfolgen, um eine Einschätzung des zu beurteilenden Bewegungsmerkmals zu treffen. Beim Bewegungsrhythmus achtet ein Beobachter bevorzugt auf die Beinbewegungen, beim Bewegungsumfang stärker auf den Wurfarm oder beim Bewegungstempo auf den Körperstamm. Hier kann ein umfangreiches Fachwissen dabei helfen, den Fokus bewusst und gezielt auf die aussagekräftigste Körperregion zu richten, wodurch das Beobachten des Bewegungsmerkmals erleichtert wird. Die Sehstrategie ist demnach relativ einfach und theoretisch ergründbar. Bei komplexen Beobachtungsaufgaben stellt sich die Sehstrategie wesentlich anspruchsvoller dar: wenn mehrere Bewegungsmerkmale gleichzeitig oder in schneller Abfolge beobachtet werden müssen, ist ein gutes Zusammenspiel zwischen Fixationen und Sakkaden erforderlich. Je besser man die Bewegung aufgrund eigener Bewegungsqualitäten kennt, desto ökonomischer kann auch das Beobachten gelingen, weil der Beurteiler die Bewegungskomplexität auf die wesentlichen Teilbewegungen und Körperpositionen reduzieren kann. Dieser Zusammenhang müsste allerdings mittels Ergebnissen von Eye-tracking-Untersuchungen überprüft werden.

6.2 Bewegungsbeurteilung und Fachwissen

Um die Tendenzen, die eben sehr allgemein aus der Betrachtung der beiden Teilindices BT_AB und BT_CDE abgeleitet wurden, differenzierter zu beschreiben, können die Grundhypothesen (siehe Kapitel 1) als Ausgangspunkt für weitere Diskussionen der Ergebnisse dienen. Bezüglich des Fachwissens (Grundhypothese 1) muss der Zusammenhang mit der Beurteilungsleistung der Schlagwurfbewegung differenziert betrachtet werden. Dass das Fachwissen positiv mit der Beurteilungsleistung der verschiedenen Beurteilungsaufgaben korreliert, kann nicht uneingeschränkt behauptet werden. Zum einen scheint es verschiedene Bereiche des Fachwissens zu geben, die eine unterschiedliche Bedeutung für das Beurteilen der Schlagwurfbewegung haben. Zum anderen hängt es von der Art der Beurteilungsaufgabe ab, wie stark die Korrelation zwischen den einzelnen Bereichen des Fachwissens und der entsprechenden Beurteilungsleistung ausfällt. Während das räumlich-zeitliche Bewegungswissen der Schlagwurfbewegung als signifikanter Prädiktor für das Beurteilen eines einzelnen Bewegungsmerkmals (Teil B des Beurteilungstests) dienen kann (siehe Kapitel 5.1.6.2), kann für das biomechanische Wissen mit der einfachsten Beurteilungsaufgabe (Teil A: Unterscheidung der Qualitätsstufen eines Bewegungsmerkmals) ein positiver Wirkungszusammenhang mit kleiner Effektstärke nachgewiesen werden (siehe Kapitel 5.1.6.3). Offensichtlich ist es vor allem bei der qualitativen Beurteilung eines einzelnen Bewegungsmerkmals (in Form einer Zuordnung von drei Bewegungsbeispielen zur jeweiligen Qualitätsstufe in Teil A oder in Form einer Zuordnung einer einzelnen Qualitätsstufe zu einem Bewegungsbeispiel wie in Teil B) hilfreich, wenn bewusstseinsfähige Wissensanteile über die Wurfbewegung vorhanden sind. Der Grund liegt vermutlich darin, dass sich ein Beurteiler nur auf ein Bewegungsmerkmal allein konzentrieren muss und somit die Aufmerksamkeit auf Aspekte lenken kann, welche die Ausführungsqualität dieses einen Bewegungsmerkmals ausmachen. Je mehr eine Person über dieses Bewegungsmerkmal weiß oder Zusammenhänge kennt, welche für die Qualität dieses Bewegungsmerkmals eine Rolle spielen, desto besser kann sie sich auf die Beobachtung dieser Details vorbereiten und gezielter analysieren. In den Bewegungsmerkmalen stecken sozusagen Fachwissensanteile, die für das inhaltliche Verstehen sowie das darauffolgende Beurteilen des jeweiligen Bewegungsmerkmals hilfreich sind. Wenn die Beurteilungsaufgabe jedoch komplexer wird, indem mehrere Bewegungsmerkmale gleichzeitig beurteilt werden müssen oder zwei Bewegungen miteinander verglichen werden müssen, so hat das bewusstseinsfähige, deklarative Fachwissen keine große Bedeutung mehr für die Qualität des Beurteilungsprozesses. Somit muss die Aussage von Girardin und Hanson (1967), dass ein detaillierteres Fachwissen zu einer besseren Einschätzung der Bewegungsqualität führt, in Frage gestellt oder zumindest um die Feststellung erweitert werden, dass es auf die Komplexität der Beurteilungsaufgabe ankommt, ob das Fachwissen einen positiven oder negativen Einfluss auf die Beurteilungsleistung hat.

Es scheint sich also bei der Beurteilung von Bewegungsmerkmalen um Prozesse zu handeln, die wenig mit dem bewusstseinsfähigen Wissen über die räumlich-zeitlichen oder biomechanischen Aspekte der Zielbewegung zu tun haben. Insofern kann der These von Radford (1990, S. 8–11) zugestimmt werden, dass deklarative Informationen mehrschichtig mit prozeduralen Wissensanteilen verknüpft werden müssen, damit sie für die Bewegungsbeobachtung nutzbar gemacht werden können. Das Ergebnis, dass 21,2 Prozent der Beurteilungsleistung von Teil B mit dem räumlich-zeitlichen Fachwissen ($\beta = 0,287$) als positiven und der Eigenrealisation ($\beta = -0,444$) als negativen Prädiktor signifikant erklärt werden können, deutet auf die Existenz eines von Tidow (1981, S. 264) behaupteten eigenrealisationsfreien Teils der prozeduralen Bewegungsvorstellung hin, der auf visuellen und verbalen Elementen beruht (oder wie Radford (1990, S. 9) es bezeichnet, als „knowledge formulated from others perform“)

und sich von einem eigenrealitätsgebundenen Teil (oder einer „knowledge formulated from learning how to perform“ (ebd.)) unterscheidet (siehe Kapitel 2.3.2). Diese eigenrealisationsfreie Bewegungsvorstellung scheint demnach für die Bewältigung einfacher Beurteilungsaufgaben auszureichen. Dies überrascht insofern, als man vermuten könnte, dass eine hohe Qualität der Eigenrealisation immer auch positiv auf die Beurteilungsfähigkeit wirkt. Es könnte sogar der Schluss gezogen werden, dass das Fachwissen nicht wie in dieser Studie in räumlich-zeitlicher und biomechanischer Hinsicht unterschieden wird, sondern in eigenrealisationsfreie und eigenrealitätsgebundene Anteile. Erstere würden dann wohl sehr stark der Definition des Fachwissens in dieser Studie entsprechen, während für zweitere eine geeignete Operationalisierung konstruiert werden müsste. Denn es müsste für die Feststellung der eigenrealitätsgebundenen Anteile überprüft werden, wie kinästhetische, akustische, vestibuläre und taktile Reafferenzen mit der prozeduralen Bewegungsvorstellung gekoppelt sind. Möglicherweise kann dann ein weiterer Prädiktor identifiziert werden, der die Beurteilungsleistung einzelner Aufgabenstellungen besser beschreibt als die in dieser Studie verwendeten Fachwissensanteile.

Wenn die Beurteilungsaufgabe sehr komplex ist, wie dies bei der synchronen Gesamtanalyse (Teil E des Beurteilungstests) beispielsweise der Fall ist, so kann das Fachwissen sogar hinderlich für eine gute Beurteilungsleistung sein, wenngleich es sich hierbei lediglich um eine These handelt, die nicht signifikant nachgewiesen werden konnte. Der standardisierte Beta-Koeffizient für das biomechanische Fachwissen bei Teil E beträgt $\beta = -0,164$ ($p = 0,205$) und für die Leistung der Bildsortieraufgabe $\beta = -0,262$ ($p = 0,045$), was in beiden Fällen für eine negative Wirkung auf die Beurteilungsleistung spricht. Testpersonen mit einem hohen biomechanischen Fachwissen und einer guten visuellen Kenntnis der Körperpositionen bei der Schlagwurfbewegung gelingt demnach die Beurteilungsaufgabe, alle Bewegungsmerkmale synchron nach einem Beobachtungsdurchgang zu beurteilen, schlechter. Dies kann damit zusammenhängen, dass Testpersonen mit einer geringen Kenntnis in diesen Bereichen die Beurteilungsaufgaben intuitiver und auf eine weniger kognitiv strukturierte Weise angehen, was letztendlich zu einer besseren Gesamteinschätzung der Bewegungsmerkmale führen kann. Diese Feststellung deckt sich mit der Aussage von Baumann (1986, S. 52f.), dass globale Bewegungsbeurteilungen, wie sie bei der Einschätzung der Meinel'schen Bewegungsmerkmale vorliegen, geringere Anforderungen an den Beurteiler stellen, weil eine ganzheitliche Repräsentation der Bewegung mit wenigen kennzeichnenden Merkmalen als kognitives Vergleichsmaterial ausreicht und kein allzu differenzierter Sollwert vorliegen muss (siehe Kapitel 2.3.2). Somit können auch Personen mit geringerem Fachwissen gute Beurteilungsleistungen erzielen. Insgesamt muss aber festgehalten werden, dass das räumlich-zeitliche und biomechanische Fachwissen keine nachweisbaren Einflussfaktoren auf die Beurteilungsleistung der komplexen Beurteilungsaufgaben der Teile C, D und E zu sein scheinen.

Hingegen ist der Bereich des **Fehlerwissens** eine lohnenswertere, wenngleich nicht sonderlich stark und vor allem nicht signifikant nachweisbare Synergiequelle für die Beurteilungsprozesse. Die positiven Wirkung auf die Beurteilung einzelner Bewegungsmerkmale ($\beta = 0,128$ ($p = 0,304$) mit Beurteilungsteil B, siehe Kapitel 5.1.6.2) und die Fehleranalyse ($\beta = 0,080$ ($p = 0,528$) mit Beurteilungsteil D, siehe Kapitel 5.1.6.1) sprechen dafür, dass die Kenntnis von Fehlerbildern das Bewegungswissen so bereichert, dass die Beurteilungsqualität zunimmt. Insofern reicht das alleinige Wissen über die ideale Bewegungsausführung als Voraussetzung für eine gute Bewegungsbeurteilung nicht aus, sondern muss um die Kenntnis von häufig auftretenden Bewegungsabweichungen bereichert werden. Somit bestätigt sich vor allem für die Beurteilungsaufgabe der Fehleranalyse die Behauptung von Radford (1990, S. 7), Biscan und Hofmann (1976) und Gangstead und Beveridge (1984), dass mit dem Wissen über Fehlerbilder die Wahrscheinlichkeit der Fehleridentifikation zunimmt. Die **visuelle Bewegungs-**

vorstellungsfähigkeit in Form von Schlüsselpositionen scheint hingegen vor allem bei komplexen Beurteilungsaufgaben negativ mit der Beurteilungsleistung zu korrelieren. Da sich diese Schlüsselpositionen am optimalen Bewegungsablauf orientieren und somit eine Ergänzung zum räumlich-zeitlichen Fachwissen darstellen, stellt die visuelle Bewegungsvorstellungsfähigkeit keine neue Qualität des Fachwissens dar, sondern weist eine starke Nähe zum semantischen Wissen über die richtige Bewegungsbeschreibung auf.

Es muss demnach die Behauptung von Neumaier (1988, S. 283), dass generell „alle dem Beurteiler verfügbaren Informationen über die zu beurteilende Bewegung“ für die Bewegungsbeurteilung eine Rolle spielen, relativiert werden, weil es nicht auf die Quantität der Einzelinformationen ankommt, die ein Beurteiler über den räumlich-zeitlichen Bewegungsverlauf sowie die biomechanischen Zusammenhänge (in Form eines abprüfbaren Faktenwissens wie im Fachwissenstest dieser Studie) weiß. Vielmehr kommt es darauf an, wie diese Vorinformationen „zur Entwicklung antizipierender Schemata im Wahrnehmungsprozess [sic!]“ (Neumaier, 1988, S. 283) genutzt werden, um auf den zweiten Teil von Neumaiers Behauptungen zurückzukommen. Dieser antizipierende Prozess scheint sich im Wissen über häufig vorkommende Bewegungsabweichungen abzubilden, denn hier richtet der Beurteiler seine Aufmerksamkeit bereits auf die Auswahl von Knotenpunkten der Bewegung, die bei späteren Beurteilungsaufgaben hilfreich für das Beobachten der Bewegung sein können. Zugleich kann eine Schlussfolgerung aus diesen Ergebnissen auch darin bestehen, dass angehende Trainer und Lehrkräfte einen Zugang zur Bewegungsbeobachtung nicht primär über die Beschreibung der idealen Bewegungsausführung in räumlich-zeitlicher Chronologie erwerben, sondern sich bereits stärker an den Knotenpunkten der Bewegung orientieren, an denen auch beobachtbare Bewegungsabweichungen auftreten können. Es kann durchaus mit Knudson (2013, S. 81f.) angenommen werden, dass sich die Fähigkeit zur Hierarchisierung der Schlüsselemente positiv auf die Bewegungsbeobachtung und Bewegungsbeurteilung auswirkt, jedoch scheint ein umfassenderes Bewegungswissen nicht einfach zu einer verbesserten Identifizierung der critical features zu führen. Vielmehr muss bei der Aneignung von Bewegungswissen bereits darauf geachtet werden, eine gewisse Hierarchie der Bewegungskriterien zu verinnerlichen, um beim Beurteilungsprozess antizipierend beobachten zu können. Hierfür scheint ein funktionales Bewegungsverständnis, wie es Göhner favorisiert, zielführender zu sein als ein räumlich-zeitliches, weil hier die Funktionalität einzelner Bewegungsteile für das Bewegungsziel bereits eine Hierarchisierung vorgibt. Es könnte also in einer Weiterführung dieser Studie durchaus interessant sein, die Aussagen von Morrison (2000, S. 22) genauer zu untersuchen, dass das verwendete Format der Bewegungsbeschreibung sowie die Art des Bewegungswissens Auswirkungen auf die Analysefähigkeit hat.

6.3 Bewegungsbeurteilung und Eigenrealisation

Der zweite wichtige Einflussfaktor, der in dieser Studie untersucht wurde, ist die Qualität der Eigenrealisation. Die Aussage von Neumaier (1988, S. 284f.), dass „die Vorstellung einer sportlichen Technik [...] ‚fragmentarisch‘ bleibt, wenn sie ausschließlich eigenrealisationsfrei erworbene Informationen umfaßt [sic!]“, wurde schon in Kapitel 6.1 diskutiert und kann indirekt bestätigt werden. Je komplexer eine Bewegungsbeurteilung wird, desto hilfreicher scheint das prozedurale Wissen zu sein, welches sich in einer hohen Demonstrationsfähigkeit (bzw. Eigenrealisation) des leichtathletischen Schlagwurfs zeigt. Wenn es darum geht, mehrere Bewegungsmerkmale gleichzeitig wahrzunehmen und einzu-

schätzen, so ist ein gut ausgeprägtes kinästhetisches Bewegungsgefühl, sozusagen als inneres Nachvollziehen des eigenen Bewegungsablaufs, gewinnbringender als ein reines Faktenwissen. Diese Erkenntnis widerspricht somit teilweise der Behauptung von Girardin und Hanson (1967), dass sowohl die Eigenrealisation als auch die theoretische Kenntnis über die Bewegung positiv mit der Beurteilungsleistung zusammenhängen.

Blickt man speziell auf die Ergebnisse von Teil D des Beurteilungstests, so kann die These von Tidow (1981, S. 264) gestützt werden, dass gerade diese Fähigkeit des Mitvollziehens einer Bewegung die Qualität der Fehleranalyse zu einem hohen Anteil mitbestimmt (siehe Kapitel 5.2.4.4). Speziell beim Identifizieren fehlerhaft ausgeführter Bewegungsmerkmale ließ sich die Qualität der Eigenrealisation als entscheidender Prädiktor signifikant ermitteln und trägt mit einem standardisierten Regressionskoeffizienten von $\beta = 0,320$ ($p = 0,008$) (im Rahmen einer hierarchischen linearen multiplen Regressionsanalyse) mit 11,7 Prozent zur Vorhersage der Beurteilungsleistung der Fehleranalyse bei. Es darf also durchaus begründet angenommen werden, dass das Identifizieren fehlerhafter Bewegungsmerkmale denjenigen Testpersonen zuverlässiger gelingt, welche die Schlagwurfbewegung besser ausführen können. Dies lässt sich möglicherweise auch mit der in Kapitel 2.3.1 genannten Analogie-Fähigkeit erklären, vom Bewegungsverhalten foveal fixierter Körperteile auf das nicht-beobachtete Verhalten anderer Körperteile schließen zu können, gerade wenn es sich um stark fehlerhafte Bewegungsmerkmale handelt. Ein Beurteiler, der die Bewegung auch als ausführender Sportler kennt, hat hier ein breiteres Repertoire an Bewegungserfahrungen zur Verfügung, mit denen er die aktuell beobachtete Bewegungsausführung vergleichen kann. Je besser dabei das eigene Bewegungskönnen ausgeprägt ist, desto deutlicher wird auch eine starke Abweichung im Bezug auf einzelne Bewegungsmerkmale wahrgenommen werden können, auch wenn nicht jedes Bewegungsmerkmal bewusst beobachtet wird. Es liegt also nahe, dass Beurteiler mit einer hohen Qualität der Eigenrealisation bei der Betrachtung der Gesamtbewegung fehlerhafte Bewegungsmerkmale relativ zuverlässig identifizieren können.

Auch auf die Frage von Neumaier (1988, S. 297), bis zu welchem Beurteilungsniveau eine in Verbindung mit der Eigenrealisation gebildete Bewegungsvorstellung ausgeglichen werden kann, kann im Rahmen dieser Studie eine vorsichtige Antwort gegeben werden. Für die Einschätzung eines einzelnen Bewegungsmerkmals mithilfe der drei Qualitätsstufen, wie sie in Teil A und Teil B des Beurteilungstests verlangt wird, reicht eine intensive, ausschließlich theoretische Schulung der Bewegungsvorstellung noch aus, vielmehr bewirkt eine hohe Qualität der Eigenrealisation bei einfachen Beurteilungsaufgaben sogar eine Verschlechterung der Beurteilungsleistung. Der Personenwert für Teil B des Beurteilungstests wird maßgeblich vom Personenwert für die Eigenrealisation bestimmt, wobei die Wirkung eindeutig negativ ist ($R^2 = 0,177$, $\beta = -0,444$, $p = 0,001$). Dies kann die Annahme stützen, dass Beurteiler mit einer hohen Eigenrealisation eine spezielle Analysefähigkeit besitzen, die sie bei komplexeren Beurteilungsaufgaben bevorteilt, bei einer einfachen Bewegungsbeobachtung aber hinderlich wirkt. Dieser Effekt könnte möglicherweise ebenfalls mit der eben genannten Analogiefähigkeit erklärt werden, denn Testpersonen mit großer eigener Bewegungserfahrung nehmen bei einer Fremdbeobachtung viele Informationen wahr, die für die Beurteilung eines einzelnen Bewegungsmerkmals überhaupt nicht nötig wären. Diese Mehraufnahme kann dazu führen, dass die eigentliche Beurteilungsaufgabe nicht mehr fokussiert durchgeführt werden kann und es zu Fehleinschätzungen kommt, die durch die unnötige, aber unbewusst ablaufende Beurteilung anderer, möglicherweise auffällig fehlerhafter Bewegungsmerkmale bedingt sind. Da es sich hierbei jedoch um extrem schwer zu messende Informationsverarbeitungsprozesse handelt, kann es sich lediglich um eine Vermutung handeln, dass die Analogiefähigkeit einfache Beurteilungsleistungen erschwert.

Bei anspruchsvolleren Beurteilungsaufgaben wie dem Vergleich von Bewegungsausführungen (Teil C), der Fehleranalyse (Teil D) oder der synchronen Gesamtbeurteilung mehrerer Bewegungsmerkmale (Teil E) ist hingegen ohne eigenrealisationsgebundene Bewegungsvorstellung eine Verschlechterung der Beurteilungsleistung zu erwarten. Diese Erkenntnis hat auch Konsequenzen für die Einschätzung und Planung eines Beurteilungsvorgangs: Wenn Bewegungen beurteilt werden sollen, bei denen der Beurteilende keine Eigenrealisationserfahrungen besitzt, so ist es ausreichend, sich auf die Qualifizierung eines einzelnen Bewegungsmerkmals zu konzentrieren. Dieser Beurteilungsvorgang ist realistisch auch ohne motorische Vorerfahrungen in der zu beurteilenden Bewegung allein durch theoretische Auseinandersetzung mit der Bewegungsausführung möglich. Wenn jedoch komplexere Beurteilungsaufgaben zuverlässig bewältigt werden sollen, kann die eigenrealisationsgebundene Bewegungsvorstellung nicht mehr kompensiert werden und der Beurteilende muss sich mit der zu beurteilenden Bewegung auch motorisch auseinandersetzen. Dabei gilt zu beachten, dass der Beurteilende die Bewegung nicht aktuell beherrschen muss, sondern auch auf frühere motorische Erfahrungen zurückgreifen kann, die bei ausreichender Übung zumeist sehr lange in stabiler Form im prozeduralen Gedächtnis verfügbar bleiben. Dabei muss die von Neumaier (1988) gestellte Frage und in der Erweiterungshypothese 3.1 formulierte Annahme offen bleiben, ab welchem motorischen Könnensstand die Beurteilungsfähigkeit nicht mehr stetig zunimmt. Ein solches Leistungsplateau, auf dem auch komplexe Beurteilungsaufgaben zuverlässig gelöst werden können, konnte im Rahmen dieser Studie nicht ermittelt werden.

Die weitere These Neumaiers (1988, S. 288), dass sich „ein gutes Eigenrealisationsniveau in der überwiegenden Zahl der Fälle auch in einem differenzierteren Wissen um die Merkmale der Idealbewegung ausdrückt“, kann anhand der Daten dieser Studie allenfalls vage vermutet werden. Bei den 59 Testpersonen deutet sich ein leicht positiver Zusammenhang zwischen dem Fachwissen und der Eigenrealisation an, was sich dahingehend interpretieren lässt, dass Testpersonen mit einem hohen Fachwissen tendenziell auch eine höhere Fähigkeit zur Eigenrealisation zeigen. Dabei ist die Korrelation zwischen der Eigenrealisation und dem räumlich-zeitlichen Bewegungswissen (Teil A des Fachwissenstests) mit $r = 0,266$ ($p = 0,042$) deutlicher ausgeprägt als zwischen der Eigenrealisation und dem biomechanischen Wissen (Teil B des Fachwissenstests) mit $r = 0,126$ ($p = 0,342$), wohingegen der Zusammenhang mit dem Fehlerwissen unbedeutend ausfällt ($r = 0,022$, $p = 0,871$). Insgesamt kann, anders als bei Girardin und Hanson (1967, S. 559), die in ihrer Untersuchung keinen Zusammenhang zwischen der Eigenrealisation und dem theoretischen Bewegungswissen feststellen konnten, die Tendenz vermutet werden, dass Eigenrealisation und Fachwissen sich gegenseitig positiv beeinflussen. Auffällig ist wiederum auch, dass als einziger Messwert die visuelle Bewegungsvorstellungsfähigkeit (gemessen mit der Sortieraufgabe der Bildreihe) mit $r = -0,149$ ($p = 0,260$) negativ mit der Eigenrealisation korreliert. Es kann demnach nicht erwartet werden, dass allein durch eine gute Beherrschung der motorischen Bewegungsfertigkeit des Schlagwurfs auch eine bewusstseinsfähige visuelle Bewegungsvorstellung über Schlüsselpositionen entsteht, die dann für Beurteilungsprozesse genutzt werden kann.

Somit kann die Grundhypothese 3, dass die Qualität der Eigenrealisation die Beurteilungsleistung durchwegs positiv beeinflusst, nicht uneingeschränkt verifiziert werden. Es trifft zwar auf komplexe Beurteilungsaufgaben zu, dass eine bessere Bewegungsdemonstrationsfähigkeit eine positive Wirkung auf die Beurteilungsleistung hat, bei einfachen Beurteilungsaufgaben ist jedoch gegenteiliges der Fall. Die Fähigkeiten, die sich ein Beurteilender durch eine gute Eigenrealisation erwirbt, müssen demnach in ihrer Wirkung auf die Beurteilungsleistung ambivalent betrachtet werden. Zusammenfassend bleibt festzuhalten, dass die beiden Einflussfaktoren Fachwissen und Eigenrealisation je nach Komplexität

der Beurteilungsaufgabe unterschiedlich wichtig sind: ein gutes Fachwissen begünstigt die Beurteilungsleistung bei einfacheren Beurteilungsaufgaben, während sich eine gute Eigenrealisation positiv auf die Beurteilungsleistung für komplexere Bewegungsanalysen auswirkt.

6.4 Bewegungsbeurteilung und Fähigkeit zum Ist-Sollwert-Vergleich

Die Fähigkeit zum Ist-Sollwert-Vergleich, welche mit dem TKBV getestet wurde, konnte nicht als wesentlicher Einflussfaktor auf die Beurteilungsfähigkeit nachgewiesen werden. Diese Feststellung ist durchaus überraschend, denn die Fähigkeit, sich eine einfache Bewegungsfolge mental vorzustellen und dann in unterschiedlicher Form abzurufen, weist eine (wenngleich sehr vereinfachte) Nähe zum Prozess der Bewegungsbeurteilung auf. Betrachtet man die Ergebnisse genauer, so spielt es einerseits eine Rolle für die Korrelationsbeziehungen, ob man bei der Testauswertung beide Teile des TKBV gemeinsam auswertet oder sie getrennt voneinander betrachtet. Die beiden Aufgabenstellungen sind doch so unterschiedlich, dass man durchaus annehmen kann, dass unterschiedliche Fähigkeiten der Bewegungsvorstellung dabei angesprochen werden. Vor allem die Gedächtnisleistung ist bei Teil 1 anspruchsvoller, weil die Endposition erst mit den angebotenen Lösungsvorschlägen verglichen werden muss, ehe die korrekte Antwort ausgewählt werden kann. Dafür erschwert bei Teil 2 die Fähigkeit, die gedanklich nachvollzogene Bewegung motorisch umzusetzen, die Lösungswahrscheinlichkeit. Bei der Regressionsanalyse von Teil C des Beurteilungstests (Bewegungsvergleich) bleiben nach Rückwärtselimination die beiden Personenwerte der Einzelteile des TKBV als plausibelste Prädiktoren übrig, zeigen aber entgegengesetzte Wirkung. Während ein hoher TKBV-2-Score auch eine Erhöhung der Beurteilungsleistung im Bewegungsvergleich vorhersagt, ist dies beim TKBV-1-Personenparameter umgekehrt (siehe Kapitel 5.1.6.4). Insofern kann durchaus hinterfragt werden, ob für die beiden Einzelteile wirklich Eindimensionalität vorliegt.

Andererseits hängt es auch von der Beurteilungsaufgabe ab, wie stark der Zusammenhang zum jeweiligen Personenfähigkeitenwert ausfällt. Es kann generell festgehalten werden, dass die Wiedererkennungsaufgabe des TKBV (Teil 1) mit den komplexeren Beurteilungsaufgaben des Bewegungsvergleichs, der Fehleranalyse und der synchronen Gesamtanalyse negativ korreliert, während sich die Beziehung zu den einfachen Beurteilungsaufgaben neutral gestaltet. Wie eben erwähnt, konnte der Personenparameter des TKBV-1 als signifikanter, negativ wirkender Prädiktor ($\beta = -0,295$, $p = 0,027$) für die Beurteilungsleistung des Bewegungsvergleichs ermittelt werden. Auch für die Beurteilungsleistung der Fehleranalyse wird er mittels Regressionsanalyse hinter der Eigenrealisation sogar als zweitwichtigster, wenngleich knapp nicht signifikanter Prädiktor ($\beta = -0,208$, $p = 0,097$) identifiziert (siehe Kapitel 5.1.6.1). Eine mögliche Erklärung könnte darin liegen, dass bei der Wiedererkennungsaufgabe neben der Fähigkeit zur mentalen Bewegungsvorstellung die Merkfähigkeit der Testperson verstärkt im Mittelpunkt steht, um die gedanklich ermittelte Endposition mit den als Lösung angebotenen Körperpositionen vergleichen zu können. Die Bewegungsrepräsentation muss also wesentlich länger im Gedächtnis bleiben und intensiver verarbeitet werden als bei Teil 2 des TKBV, wo die Endposition sofort nach der letzten Bewegungsanweisung eingenommen werden kann. Testpersonen, die sich mit dieser verhältnismäßig langsam ablaufenden Entscheidungsaufgabe von Teil 1 leichter tun, sind möglicherweise bei den komplexen Beurteilungsaufgaben mit der sofort zu treffenden Entscheidung der Qualitätseinschätzung von Bewegungsmerkmalen überfordert.

Teil 2 des TKBV forciert hingegen schnell ablaufende Prozesse, die mentale Bewegungsvorstellung in eine motorische Aktion zu überführen. Diese Fähigkeit ist zumindest nicht kontraproduktiv für die komplexeren Beurteilungsaufgaben und mit der Beurteilungsleistung des Bewegungsvergleichs sogar eindeutig positiv korreliert. Der Personenparameter des TKBV-Teil2 konnte immerhin als zweitwichtigster, positiv wirkender Prädiktor ($\beta_{\text{TKBV2}} = 0,214$, $p = 0,106$) auf die Beurteilungsleistung des Bewegungsvergleichs nachgewiesen werden, auch wenn die 0,05-Signifikanzgrenze leicht überschritten wurde. Das Grundmuster der Beantwortung bei Teil 2 des TKBV, ohne vorgegebene Antwortmöglichkeiten entscheiden zu müssen, entspricht im Gegensatz zu Teil 1 eher den Anforderungen der in dieser Studie verwendeten Beurteilungsaufgaben, so dass damit möglicherweise der Unterschied zwischen beiden Teilen des TKBV erklärt werden kann.

Insgesamt muss aber festgehalten werden, dass der TKBV nicht den erwarteten Beitrag zur Aufklärung der Einflussfaktoren auf die Beurteilungsfähigkeit geleistet hat. Zugleich kann dieses Ergebnis als Bestätigung für die Aussagen von Knudson (2013, S. 77) dienen, dass die Voraussetzungen für eine gute Bewegungsbeurteilung auch bewegungsspezifisch ausgeprägt sein müssen, d. h. trotz einer guten allgemeinen Bewegungsvorstellungsfähigkeit kommt es entscheidend darauf an, sich den Bewegungsablauf des Schlagwurfs mental richtig vorstellen zu können (siehe Kapitel 2.3.2.1). Eine allgemein geschulte Bewegungsvorstellungsfähigkeit kann die spezielle Beschäftigung mit der zu beobachtenden Bewegungen allenfalls bereichern, keinesfalls aber ersetzen. Somit kann die Grundhypothese 2 nicht bestätigt werden, dass von einer guten allgemeinen Bewegungsvorstellungsfähigkeit eine positive Wirkung auf die Beurteilungsleistung des Schlagwurfs ausgehen würde.

6.5 Bewegungsbeurteilung und Beurteilungserfahrung

Grundsätzlich lässt sich die These nicht generell pauschalisieren, dass eine größere Beurteilungserfahrung automatisch zu besseren Beurteilungsleistungen führt. Vielmehr muss die Art der Beurteilungserfahrung ebenso wie die Komplexität der Beurteilungsaufgabe in die Schlussfolgerungen miteinbezogen werden. Es überrascht, dass die Testpersonen ohne Beurteilungsvorerfahrungen bei komplexeren Beurteilungsaufgaben überdurchschnittlich gut abschneiden. Der Erfahrungsvorsprung der Beurteiler, die bereits leichtathletische Schlagwurfbewegungen analysiert und bewertet haben, tritt vor allem bei der Fehleridentifikation (Teil D) gegenüber den Gruppen mit weniger spezifischen Beurteilungserfahrungen zu Tage. Dies spricht dafür, dass die Fehleridentifikation eine besonders typische Aufgabe im Alltag von (angehenden) Sportlehrkräften und Trainern darstellt, bei der sich der Erfahrungsvorsprung besonders deutlich auswirkt. Die anderen Beurteilungsaufgaben wie das Beurteilen eines einzelnen Bewegungsmerkmals oder der Bewegungsvergleich liefern keine so eindeutigen Ergebnisse und hängen somit auch nicht speziell von der Beurteilungserfahrung ab. Die synchrone Gesamtanalyse der sechs Bewegungsmerkmale ist hingegen eine so komplexe Beurteilungsaufgabe, dass die Beurteilungserfahrung auch keinen Vorteil für das richtige Einschätzen der Bewegungsmerkmale darstellt.

Ein nicht zu erwartendes Ergebnis liegt auch darin, dass die Personen, welche bereits Beurteilungserfahrungen in Sportarten mit dem leichtathletischen Schlagwurf ähnlichen Bewegungen (z. B. Handball, Volleyball) hatten, den gesamten Beurteilungstest betreffend eine schlechtere Beurteilungsleistung als völlig beurteilungsunerfahrenen Testpersonen erzielten. Eine mögliche Erklärung könnte sein, dass diese spezielle Gruppe durch ihre bereits vorhandene Beurteilungserfahrung von schlag(wurf)ähnlichen Bewegungen ein sportartspezifisches Beurteilungsmuster verinnerlicht hat, das trotz allgemeiner

Übereinstimmungen mit der leichtathletischen Schlagwurfbewegung in einigen konkreten Beurteilungspunkten davon abweicht. Die von anderen Schlagbewegungen beeinflussten Testpersonen sind nicht in der Lage, diese Unterscheidung bewegungsspezifisch zu vollziehen. Diese vorgeprägte Personengruppe ist dann nicht fähig, die Beurteilungskriterien, welche im Rahmen des Beurteilungstests für die leichtathletische Schlagwurfbewegung vermittelt werden, in das bereits bestehende Beurteilungsmuster der Fremdsportart zu integrieren. Möglicherweise entsteht durch die speziellen leichtathletikspezifischen Bewegungskriterien, die nicht mit den bereits bekannten Schlag(wurf)bewegungen übereinstimmen, eine Verunsicherung des Beurteilungsvermögens, welche sich bei der Bearbeitung der Beurteilungsaufgaben negativ auswirkt. Bedenkt man, dass die Beurteilungserfahrungen dieser Gruppe hauptsächlich in den Sportarten Handball und Volleyball gemacht wurden, so erscheint dies nachvollziehbar, denn ein Schmetterschlag im Volleyball oder ein Schlagwurf im Handball wird trotz vieler Ähnlichkeiten zum leichtathletischen Schlagwurf in einigen Bewegungsdetails anders ausgeführt, was wiederum zu Fehleinschätzungen bei der Qualifizierung der im Beurteilungstest vorgegebenen Wurfbewegungen führen kann. Beobachter ohne Beurteilungserfahrung haben hingegen weder spezifische Beurteilungserfahrungen von Schlag(wurf)bewegungen noch von anderen sportlichen Bewegungen, so dass die vorgegebenen Beurteilungskriterien in keiner Konkurrenz zu bereits bekannten Beurteilungsmustern stehen und somit irritationslos angewendet werden können.

Die Grundhypothese 4, dass eine hohe Spezifität der Beurteilungsvorerfahrung zu einer besseren Beurteilungsleistung führt, kann somit nicht generell bestätigt werden, sondern trifft speziell für den Prozess der Fehleridentifikation zu.

6.6 Praxisrelevanz und Kritik der Ergebnisse

Am Ende der Diskussion soll nun die die Praxisrelevanz der Ergebnisse herausgestellt werden. Zu Beginn der Studie wurde ein Forschungsdefizit bezüglich der Frage identifiziert, welche Variablen Einfluss auf die Beurteilung sportlicher Bewegungen nehmen. Die Ergebnisse machen deutlich, dass es von der Beurteilungsaufgabe abhängt, welche Zusammenhänge mit möglichen Einflussfaktoren behauptet werden können. Es ist also nicht davon auszugehen, dass es eine prädestinierte Konstitution eines Beurteilers gibt, welche für die Beurteilung sportlicher Bewegungen generell besonders günstig ist, sondern es hängt von der konkreten Beurteilungssituation ab, welche Fähigkeiten ein Beurteiler besitzen sollte, um erfolgreich beurteilen zu können. Dabei lässt sich eine Zerteilung der in dieser Studie entwickelten Beurteilungsaufgaben vermuten, die als Gradmesser für die grundsätzliche Unterscheidung von Beurteilungssituationen dienen könnte. Demzufolge macht es einen entscheidenden Unterschied, ob nur ein einzelnes oder mehrere Bewegungsmerkmale gleichzeitig beurteilt werden müssen. Daraus lassen sich erste methodische Grundsätze für die Entwicklung der Beurteilungsfähigkeit sportlicher Bewegungen ableiten: gemäß der Maxime ‚Vom Einfachen zum Schweren‘ sollte demnach erst die Qualifizierung der einzelnen Beurteilungsmerkmale geübt und beherrscht werden, ehe komplexere Beurteilungsaufgaben mit mehreren Qualitätsentscheidungen verlangt werden.

Für den Beurteilungsprozess hat sich sowohl die Orientierung an den Meinel’schen Bewegungsmerkmalen als auch deren dreistufige Qualifizierung, eine Unterscheidung der Bewegungsausführung als ideal ausgeführt (Qualitätsstufe 1) oder gering (QS2) bzw. stark (QS3) von der Idealausführung abweichend zu treffen, als praktikabel erwiesen. Diese übergeordneten Bewegungsmerkmale müssen aber für jede Bewegung spezifisch beschrieben werden, damit ein Beurteiler vor der Bewegungsbeobachtung weiß, auf welche Beobachtungsdetails er sich konzentrieren muss, bevor er das entsprechende

Bewegungsmerkmal einschätzt, wie dies im Beurteilungsraster dieser Studie (siehe Appendix Tab. 53) versucht wurde. So muss der Beurteiler leichtathletischer Wurfbewegungen beispielsweise wissen, dass er zur Einschätzung des Bewegungsumfangs des Schlagwurfs neben der Oberkörperverwringung vor dem Abwurf auf die Wurfarmstreckung achten muss, während er beim Kugelstoß stattdessen die Beugung des Druckbeins mitberücksichtigen muss. Es kommt also bei der Beschreibung der Bewegungsmerkmale maßgeblich darauf an, nur die relevanten Technikdetails der zu beurteilenden Bewegung zu einem Beobachtungscluster zusammenzufassen, welches dann vom Beurteiler kategorisch eingeschätzt wird. Hierbei kann es zwar auch zu synchronoptischen Wahrnehmungsprozessen kommen, wenn beispielsweise bei der Bewegungskopplung gleichzeitig ablaufende Körperteilbewegungen beurteilt werden müssen, jedoch erleichtert zumindest die Zusammenfassung mehrerer Technikdetails zu einem Bewegungsmerkmal dessen Beurteilung. Die Bewegungsharmonie stellt dabei das allgemeinste Merkmal dar, welches noch ohne konkrete technische Beobachtungsdetails auskommt und sich daher besonders für ungeübte Beurteiler eignet, einen ersten Zugang zur Bewegungsbeobachtung zu gewinnen, wenngleich auch hierfür schon eine allgemeine Vorstellung vom Bewegungsverlauf vorliegen sollte. Daher erscheint es sinnvoll, dass Bewegungsbeurteilungen auf der Grundlage einer strukturierten Bewegungsvorstellung der zu beobachtenden Technik erfolgen, welche vor allem bei ungeübten Beurteilern frühzeitig ausgebildet werden sollte.

Wenn man die Trainierbarkeit der Beurteilungsfähigkeit in den Blick nimmt, so muss grundsätzlich bedacht werden, dass es für Beurteilungsanfänger zunächst ausreichend ist, sich auf ein einzelnes Bewegungsmerkmal zu konzentrieren und dieses qualifiziert einzuschätzen. Als Vorbereitung auf diese Beurteilungsaufgabe kann die eben erwähnte Bewegungsvorstellung über ein fundiertes räumlich-zeitliches Fachwissen über den Bewegungsverlauf, welches durchaus noch in eigenrealisationsfreier Form erworben werden kann, gewonnen werden. Auch die sinnesphysiologischen Erkenntnisse der Bewegungsbeobachtung (siehe Kapitel 2.2.1) unterstützen diese Einschätzung, dass ein einzelnes Bewegungsmerkmal noch aufmerksam wahrgenommen werden kann, wenn dem Beurteiler bewusst ist, auf welche Körperregion er seinen Blick richten muss. Neben dem Wissen über den Bewegungsverlauf scheint es daher auch vorteilhaft zu sein, wenn der Beurteiler eine Beobachtungsstrategie ausbildet, mit der er das Bewegungsmerkmal möglichst effektiv einschätzen kann, wie dies den Testpersonen im Rahmen des Beurteilungstests mittels Lehrvideos vermittelt wurde. Denn die Fokussierung auf das richtige Körperteil zum richtigen Zeitpunkt der Bewegung stellt eine entscheidende Anforderung dar, um die nötigen Informationen über die Qualität der Bewegungsausführung visuell aufnehmen zu können. Es ist zwar unmöglich, zwei voneinander entfernte Details, die möglicherweise für die Einschätzung eines Bewegungsmerkmals gleichzeitig beobachtet werden sollen, fokussiert zu betrachten, aber die bereits erwähnte Zusammenfassung zu einem Beobachtungscluster erleichtert die Ausbildung effektiver Blickstrategien (beispielsweise in der Mitte zwischen den beiden Details), um eine qualifizierte Einschätzung treffen zu können.

Beurteiler mit einer guten Eigenrealisation brauchen für einfache Beurteilungsaufgaben offensichtlich eine andere Beobachtungsstrategie, denn sie erzielen hierbei deutlich schlechtere Beurteilungsergebnisse. Dies könnte daran liegen, dass die Sicherheit im motorischen Bewegungsvollzug zur Unterschätzung der Beurteilungsaufgabe führt. Wahrscheinlicher ist aber, dass gute Eigenrealisierer Fremdbewegungen kritischer beurteilen, wenn als Referenzpunkt die Ausführungsqualität der eigenen Bewegungsausführung dient. Da ihr Fachwissen über den Bewegungsverlauf viele eigenrealisationsgebundene Anteile enthält und somit eine stärker mit kinästhetischen Empfindungen angereicherte Bewegungsvorstellung vorliegt, kann es bei der Einschätzung einzelner Bewegungsmerkmale durchaus zu Beurteilungsverzerrungen kommen. Daher müssen dieser Gruppe möglicherweise die Qualitätsstufen

der einzelnen Bewegungsausführungen anhand der Zielgruppe noch stärker bewusst gemacht werden, um den Adressatenbezug der Beurteilung (siehe Kapitel 2.3.2.4) realistisch einzuschätzen.

Bei komplizierteren Beurteilungssituationen, in denen mehrere Bewegungsmerkmale gleichzeitig beobachtet werden müssen, ist eine höhere Qualität der Eigenrealisation offensichtlich von Vorteil, wodurch die sinnesphysiologische Überforderung der synchronoptischen Wahrnehmung offensichtlich teilweise kompensiert werden kann. Insofern sollte, bevor einem Beurteiler solche anspruchsvollen Beurteilungsaufgaben wie die Fehleranalyse zugemutet werden, eine eigenrealitätsgebundene Bewegungsvorstellung vorliegen. Der Beurteiler kann diese mittels eigener Bewegungserfahrung erwerben, indem er die zu beurteilenden Bewegungen selbst motorisch erlernt und sich mit ihnen bis zu einer gewissen Könnensstufe auseinandersetzt. Dies rechtfertigt die sportpraktische Anforderung an Trainer und Sportlehrkräfte, selbst eine gewisse Demonstrationsfähigkeit der zu vermittelnden Bewegungen vorweisen zu können, damit die Bewegungsbeurteilung bei der Betreuung von Athletinnen und Athleten oder Schülerinnen und Schülern gelingt. Zugleich lohnt es sich, auch über Möglichkeiten nachzudenken, diese eigenrealisationsgebundenen Anteile nicht nur per se durch eigenen Bewegungsvollzug zu erwerben, sondern in Form eines mentalen Trainings auszubilden, weil auf diese Weise ein bewussterer Zugang zur kinästhetischen Bewegungsvorstellung gelingen könnte. Dies stellt eine interessante Forschungsfrage für eine Weiterführung dieser Studie dar, ob sich die Beurteilungsleistung durch ein derartiges mentales Training verbessern lässt. Vorschläge und methodische Anregungen dazu gibt es schon, untersucht wurden diese Möglichkeiten aber bisher hauptsächlich im Bezug auf eine Unterstützung des motorischen Lernens. Kritisch ist bei diesen Schlussfolgerungen zu erwähnen, dass es sehr vereinfacht gedacht ist, von einer guten Eigenrealisation auf das eigenrealisationsgebundene Fachwissen zu schließen. Es müsste vielmehr darüber nachgedacht werden, wie dieser eigenrealisationsgebundene Anteil methodisch überprüft werden kann, um genauere Aussagen über die Zusammenhänge mit der Beurteilungsleistung anstellen zu können. Die Messung der Eigenrealisation selbst, wie sie in dieser Studie durchgeführt wurde, kann diesbezüglich nur eine vage Annäherung sein.

Von den fünf Beurteilungsaufgaben stellt die synchrone Gesamtanalyse (Teil E) die höchsten Anforderungen an die Testpersonen. Wie schon erwähnt, ist das optische System sinnesphysiologisch gar nicht in der Lage, mehrere Beurteilungskriterien gleichzeitig zu beobachten und qualitativ einzuschätzen. Die Testergebnisse bestätigen diese Einschätzung zum einen durch den geringsten durchschnittlichen Personenparameter, der bei den Beurteilungsaufgaben erreicht wurde, zum anderen durch die Tatsache, dass keine signifikanten Prädiktoren für die Beurteilungsleistung der synchronen Gesamtanalyse gefunden werden konnten. Dies lässt den Schluss zu, dass es für die Beurteilenden eine Überforderung darstellt, die gesamte Bewegung zu beobachten und anhand von sechs Bewegungsmerkmalen differenzierte Aussagen über die Ausführungsqualität der Bewegung zu treffen. Dies ist aber genau die Anforderung, die beispielsweise beim Benoten einer Technikdemonstration durch Schülerinnen und Schüler geleistet werden muss, nämlich die Gesamtbewegung bei nur einem Beobachtungsdurchgang zu benoten und gleichzeitig noch inhaltlich feststellen zu können, welche Bewegungsmerkmale gut oder schlecht ausgeführt waren. Insofern legt diese Studie den Schluss nahe, dass es zwar durchaus möglich ist, den Gesamteindruck (im Sinne der Bewegungsharmonie) einer Bewegung einigermaßen zuverlässig einzuschätzen, ohne jedoch detailliert begründen zu können, warum man zu dieser Einschätzung kam. Wenn man diesen hohen fachdidaktischen Anspruch an eine Bewegungsbeurteilung hat, auch noch Rückmeldung über die Qualität der Bewegungsmerkmale geben zu können, sollte man über die Unterstützung durch technische Hilfsmittel nachdenken, um beispielsweise die zu beurteilende Bewegung mehrmals zu wiederholen oder mit Zeitlupe und Standbildern genauer zu analysieren. Alternativ kann die Beurteilung auch anhand mehrerer Beobachtungsdurchgänge erfolgen, indem der Sportler

die Bewegung mehrmals hintereinander vollzieht, wie dies einige Autoren ohnehin empfehlen, um eine vollständige Bewegungsanalyse überhaupt realistisch durchführen zu können (siehe Kapitel 2.2.2.4). Hier besteht auch wieder ein Fortführungspotential dieser Studie, indem überprüft wird, ob sich durch die Beobachtungswiederholung eine Verbesserung der Beurteilungsleistung einstellt. Daten zu dieser Fragestellung wurden im Rahmen dieser Studie schon erhoben, müssen aber erst ausgewertet und mit den Ergebnissen der einfachen Bewegungsbeobachtung in Beziehung gesetzt werden, um genauere Aussagen dazu treffen zu können.

Die Beurteilungsaufgabe des Vergleichs eines Bewegungsmerkmals bei zweimaliger Bewegungsausführung durch den gleichen Sportler (Teil C des Beurteilungstests) stellte ebenfalls eine sehr herausfordernde Aufgabe für die Testgruppe dar, die nochmals andere Fähigkeiten erfordert als beispielsweise die Fehleranalyse und somit als eigener Aufgabentyp der Bewegungsbeurteilung registriert werden muss. Vor allem an die Merkfähigkeit werden sehr hohe Ansprüche gestellt, weil die Qualität der Bewegungsmerkmale nach der ersten Bewegungsausführung solange memoriert werden muss, bis die Beobachtung der zweiten Bewegungsausführung abgeschlossen ist, und erst dann ein Vergleich vorgenommen werden kann. Interessanter Weise besteht gerade in der Ähnlichkeit dieser Beurteilungsaufgabe mit Teil 1 des TKBV, eine Bewegungsfolge mental nachzuvollziehen und dann die vorgestellte Endstellung mit fünf Körperpositionen zu vergleichen, ein mögliches Potential zur Trainierbarkeit der Fähigkeit zum Bewegungsvergleich. Zumindest erweist sich der TKBV-Teil1 als signifikanter Prädiktor für die Beurteilungsleistung von Teil C und es könnte mit ähnlichen Aufgaben, wie sie im Exkurs dieser Studie im Rahmen des Bewegungsbeobachtungstests (siehe Kapitel 4.3.3) bereits entwickelt wurden, die allgemeine mentale Fähigkeit zum Vergleich von Körperteilbewegungen geschult werden. Ob dieses spezielle mentale Training dann Auswirkungen auf die Beurteilungsleistung bei Aufgabenstellungen wie in Teil C dieser Studie hat, könnte dann wiederum statistisch untersucht werden. Jedenfalls kann der Vergleich von Bewegungsmerkmalen bei unterschiedlichen Bewegungsausführungen als eigenständige Beurteilungsfähigkeit gelten, die auch bei der methodischen Ausbildung der Bewegungsbeurteilung gesondert geschult werden muss.

Über die Beurteilungserfahrung als letzter möglicher Einflussfaktor können im Rahmen dieser Studie nur bedingt Aussagen getroffen werden. Als signifikanter Prädiktor konnte sie jedenfalls für keine der Beurteilungsaufgaben nachgewiesen werden, was möglicherweise auch daran liegen könnte, dass die Einteilung der Testpersonen in fünf Vorerfahrungsstufen zu differenziert war. Die inhaltlich begründete Einteilung der Probandengruppe in fünf Vorerfahrungsstufen lässt nach einem nicht-parametrischen Gruppenvergleich jedenfalls nur vorsichtige Vermutungen zu. Es scheint jedenfalls die Annahme bestätigt, dass die Beurteilungssituation der Fehleranalyse (Teil D) umso besser gelöst werden kann, je größer die Erfahrung beim Beurteilen sportlicher Bewegungen im Allgemeinen und der leichtathletischen Wurfbewegung im Speziellen ist. Dies kann möglicherweise daran liegen, dass das Identifizieren fehlerhafter Bewegungsmerkmale eine im Unterrichts- und Trainingsalltag häufig vorkommende Aufgabe darstellt, die beispielsweise bei der Bewegungsrückmeldung permanent praktiziert wird und daher mit spezifischer Beurteilungserfahrung umso besser gelöst werden kann. Bei der Konstruktion eines Trainingsprogramms zur Verbesserung der Beurteilungsleistung empfiehlt es sich daher, beim Üben solcher Beurteilungssituationen für Beurteilungserfahrene möglicherweise anspruchsvollere Aufgaben oder Bewegungsbeispiele anzubieten als für Unerfahrene. Als problematisch stellte sich auch die Gruppe mit Vorerfahrungen beim Beurteilen schlagwurfähnlicher Bewegungen heraus, die durch unterdurchschnittlich schlechte Beurteilungsleistungen auffiel. Es kann demnach durchaus zu verzerrten Beurteilungsprozessen kommen, wenn man die zu beurteilende Bewegung aus einer ande-

ren Sportart kennt und diese zwar strukturell ähnlich ausgeführt wird, sich jedoch in einzelnen Bewegungsdetails oder in der Zielsetzung der Bewegung (z. B. Präzision statt Weite) unterscheidet. Hier muss vor allem versucht werden, die vorhandene Spezifität der Bewegungsvorstellung in Richtung der zu beurteilenden Bewegung zu verändern, indem die Unterschiede in der Bewegungskonzeption bewusstgemacht werden. Ähnlichkeiten zum Umlernen motorischer Bewegungen deuten sich hierbei an, vor allem, wenn man bedenkt, dass Testpersonen, die noch keinerlei Beurteilungsvorerfahrung haben, bessere Beurteilungsergebnisse erzielen, weil sie ähnlich wie beim Bewegungsneulernen ohne irritierende Vorerfahrungen an die Beurteilungsaufgabe herangehen. Insofern lässt sich durchaus begründet annehmen, dass eine Bewegungsbeurteilung je nach Beurteilungssituation spezifisch erlernt werden muss und die Vorerfahrungen schon sehr genau mit der zu beurteilenden Bewegung übereinstimmen müssen, um die Beurteilungsleistung positiv zu beeinflussen.

Abschließend bleibt zu erwähnen, dass die angewendeten Testverfahren durchaus kritisch betrachtet werden müssen, da es abgesehen vom Test zur Kontrolle der Bewegungsvorstellung keine wissenschaftlichen Vorarbeiten gibt, an denen sich diese Studie hätte orientieren können. Trotzdem lassen die Ergebnisse einige erwartbare, aber auch viele überraschende Interpretationen zu, die in weiteren Untersuchungen überprüft werden könnten. Speziell an den Einzelaufgaben des Beurteilungstests kann nach dieser ersten Erprobung weitergearbeitet werden, um die Messbarkeit der Beurteilungsleistung noch zu verbessern. Die Aufgabe der Fehleranalyse (Teil D) besitzt dabei das höchste fachdidaktische Potential, weil das Identifizieren und die Rückmeldung über mangelhaft ausgeführte Bewegungsmerkmale zur Kernaufgabe von Lehrpersonen beim Betreuen motorischer Lernprozesse gehört. Die Hoffnung, erste Erkenntnisse über mögliche Einflussfaktoren auf die Beurteilungsleistung zu identifizieren und das Konstrukt der Bewegungsbeurteilung überhaupt messbar zu machen, hat sich dabei durchaus erfüllt.

7 Ausblick

In der Einleitung dieser Arbeit wurde betont, wie komplex sich die Wurfbewegung darstellt und welche anatomischen Besonderheiten den Menschen auszeichnen, dass er sich überhaupt als guter Werfer in der Evolution behaupten konnte. Im Laufe dieser Studie hat sich auch gezeigt, dass die Beurteilung der Wurfbewegung ein sehr komplizierter Prozess ist, in dem viele mentale Fähigkeiten zusammenspielen. Um überhaupt einen Wurf beobachten zu können, kommt das visuelle System schon an seine sinnesphysiologischen Grenzen. Eine gute Beobachtungsstrategie ist notwendig, um Informationen über die entscheidenden Teilbewegungen zu speichern und sie mit der Bewegungsvorstellung eines ideal ausgeführten Schlagwurfs zu vergleichen. Und letztlich müssen die beobachteten Bewegungsmerkmale dann auch noch derart beurteilt werden, dass differenziert werden kann, welche Teile einer Bewegung schon zufriedenstellend ausgeführt wurden und welche noch defizitär sind. Dass dies ein anspruchsvoller Prozess ist, der nur schwer zu operationalisieren sein wird, wohnt dem Phänomen der Bewegungsbeurteilung inne.

Trotz der zu erwartenden Schwierigkeiten wurde das Forschungsanliegen formuliert, mögliche Einflussfaktoren auf die Beurteilungsleistung des Schlagwurfs zu identifizieren und in ihrer Bedeutung zu untersuchen, um die gewonnenen Erkenntnisse im Rahmen der Lehramts- oder Trainerausbildung direkt für die Vermittlung der Beurteilungsfähigkeit sportlicher Bewegungen zu nutzen. Die wichtigste theoretische Erkenntnis dieser Studie liegt darin, dass man beim Beurteilen sportlicher Bewegungen nicht von einem einheitlichen Konstrukt sprechen kann, sondern vielmehr berücksichtigen muss, worin die Beurteilungsaufgabe besteht. Die fünf in dieser Studie entwickelten Beurteilungssituationen können ein erster Anhaltspunkt für eine theoretische Differenzierung des Begriffs der Bewegungsbeurteilung sein, der aber noch weiter überdacht werden muss, um in zukünftigen Studien damit arbeiten zu können. Schließlich zeigen die Ergebnisse, dass das Beurteilen von Bewegungen ein sehr spezifischer Prozess ist, der nicht einfach verallgemeinert oder in einem gemeinsamen Beurteilungsindex ausgedrückt werden kann.

Dass sich die Eigenrealisation als wichtiger Faktor für die Beurteilungsleistung erwies, bestätigt die Annahme, dass man eine Bewegung nicht nur äußerlich beschreiben, sondern sie innerlich nachvollziehen können muss, um sie umfassend beurteilen zu können. Je komplexer eine Beurteilungsaufgabe wird, desto positiver wirkt sich die Qualität der Eigenrealisation auf die Beurteilungsleistung aus. In einfacheren Beurteilungssituationen hingegen verschlechtert eine gute Bewegungsdemonstrationsfähigkeit die Beurteilungsleistung, was sicherlich ein überraschendes Ergebnis dieser Studie darstellt. Dies kann aber teilweise durch ein gut ausgeprägtes räumlich-zeitliches Fachwissen kompensiert werden. Es stellt sich die Frage, in welcher Form die Eigenrealisation überhaupt auf die Beurteilungsfähigkeit wirkt. Möglicherweise wird das zunächst eigenrealisationsfreie Fachwissen, wie es durch Bewegungsbeschreibungen der Fachliteratur erworben werden kann, mit zunehmendem Bewegungskönnen um eigenrealisationsgebundene Anteile ergänzt, welche einen Beurteiler dann auch zu einer differenzierteren Bewegungsbeobachtung befähigen. Um diesen Zusammenhang zu überprüfen, müsste überlegt werden, wie das Vorhandensein eigenrealisationsgebundener Anteile im Fachwissen überprüft werden kann. Hier müssten vor allem die kinästhetischen, taktilen und vestibulären Formen der Bewegungswahrnehmung (siehe Abb. 2), welche in dieser Studie bewusst ausgespart blieben, auch im Bezug auf die Beurteilung der Eigenbewegung berücksichtigt werden. Eine mögliche neue Hypothese könnte

dahingehend formuliert werden, dass die visuelle Beurteilung von Fremdbewegungen durch eine verbesserte kinästhetische Bewegungswahrnehmungsfähigkeit positiv beeinflusst wird.

Außerdem ermöglicht diese Studie einen Ausblick auf die Trainierbarkeit der Beurteilungsfähigkeit sportlicher Bewegungen. Dass Testpersonen, welche bereits leichtathletische Schlagwurfbewegungen beurteilt haben, die komplexe Beurteilungssituation der Fehleridentifikation besser bewältigen als Beurteilungsunerfahrene, spricht für den positiven Einfluss des Erfahrungsvorsprungs. Es könnte nun in der Fortführung dieser Studie überlegt werden, auf welche Weise die Beurteilungsfähigkeit am Besten trainiert werden kann. Ein entscheidender Schlüssel scheint in der Vorstellungsqualität der zu beurteilenden Bewegung zu liegen, welche über ein gezieltes mentales Training positiv manipuliert werden könnte. Vor allem die eben angesprochene Verbindung des räumlich-zeitlichen Fachwissens mit kinästhetischen Bewegungseindrücken könnte hierdurch zu verbessern versucht werden. Es ließe sich daraus wiederum ein interessantes Forschungsanliegen ableiten, ob durch mentales Training die Beurteilungsleistung bei einfachen Beurteilungsaufgaben verbessert werden kann. Eine andere, denkbare Perspektive der Anwendbarkeit der Methodik dieser Studie liegt in der Frage, ob sich die mittels medial präsentierter Bewegungsbeispiele gemessene Beurteilungsleistung tatsächlich mit der Beurteilungsleistung in der Realsituation deckt. Noch erkenntnisreicher wäre eine Untersuchung darüber, ob sich mit dem Üben mittels medial präsentierter Bewegungsbeispiele die Beurteilungsleistung in der Realsituation verbessern lässt. Beurteilungsaufgaben des in dieser Studie entwickelten Beurteilungstests könnten dann als Trainingsinstrumentarium dienen, um Beurteilungsvorerfahrung aufzubauen und dann in echten Beurteilungssituationen anzuwenden.

In dieser Studie wurden nur Testpersonen untersucht, bei denen bei der Ausführung der Schlagwurfbewegung noch keine motorische Einschränkung aufgrund fortgeschrittenen Alters zu erwarten war. Da die Schlagbewegung in verschiedenen Sportarten im Rahmen des Sportstudiums regelmäßig angewendet wird, kann man davon ausgehen, dass im Erhebungszeitraum auch die bestmögliche Bewegungsqualität der Testpersonen gemessen wurde. Anders sieht es beispielsweise bei älteren Sportlehrkräften aus, bei denen die Fähigkeit zur bestmöglichen Bewegungsdemonstration schon viele Jahre zurückliegt. Reicht es aus, dass die Bewegung im Laufe der aktiven Sportkarriere einmal gut beherrscht wurde, oder müssen sich Sportlehrkräfte fortwährend darum bemühen, sich durch Training eine hohe Demonstrationsfähigkeit sportlicher Bewegungen zu erhalten, weil davon auch die Beurteilungsleistung abhängt? Abgesehen von der Schwierigkeit, dass diese vergangene Fähigkeit zur Bewegungsdemonstration nicht leicht rekonstruiert werden kann, ergibt sich bei älteren Personen ein völlig anderes Konstellationsgefüge, das nicht mit dem der aktuell in dieser Studie untersuchten Probandinnen und Probanden übereinstimmt. Insofern können die Ergebnisse dieser Studie hilfreich sein für die Ausbildung angehender Sportlehrkräfte, eine Übertragung auf Sportlehrerinnen und Sportlehrer mit mehr Beurteilungserfahrung, aber einem von der früheren bestmöglichen Bewegungsausführung abweichenden Eigenrealisationsniveau erscheint möglich, würde aber einiger methodischer Umstrukturierungen bedürfen.

Auch die in dieser Studie nicht berücksichtigten Einflussgrößen der Bewegungsbeobachtung könnten ein Ausgangspunkt für fortführende Einzeluntersuchungen sein. Messmethoden des visuell-dynamischen Sehvermögens müssten getestet und auf die sportliche Beobachtungssituation angewendet werden, um zu überprüfen, ob es sich um einen leistungslimitierenden Faktor der Bewegungsbeurteilung handeln könnte. Ebenso könnte die Beobachtungsstrategie beim Beurteilen der Schlagwurfbewegung mittels Eye-tracking ermittelt werden, um beispielsweise herauszufinden, ob beurteilungserfahrene Testpersonen andere Körperregionen fixieren als Beurteilungsunerfahrene. Dies könnte wiederum für

die Trainierbarkeit der Bewegungsbeurteilung gewinnbringend sein, weil dann auch konkrete Beobachtungsstrategien erlernt werden könnten, um mit der Komplexität der Informationsaufnahme bei synchron-optischen Wahrnehmungsvorgängen, die für die Bewegungsbeurteilung eine zentrale Rolle spielt, besser zurecht zu kommen. Es könnte aber auch schon das Wissen eines Beurteilers über eine sinnvolle Beobachtungsstrategie entscheidend für eine effektive Bewegungsbeobachtung sein. Dies ließe sich beispielsweise verhältnismäßig einfach ermitteln, indem gefragt wird, auf welche Körperregion(en) eine Testperson bei der Beurteilung eines vorgegebenen Bewegungsmerkmals den Fokus richtet.

Abschließend lässt sich zusammenfassen, dass diese Studie als erster Versuch gelten kann, die Beurteilungsfähigkeit sportlicher Bewegungen messbar zu machen und mögliche Einflussfaktoren auf die Beurteilungsleistung zu identifizieren. Die Testverfahren können als Ausgangspunkt für weitere methodische Überlegungen und Modifikationen dienen, um den Erkenntnisgewinn im Bereich der Bewegungsbeurteilung zu erweitern. Die Ergebnisse müssen zwar entsprechend vorsichtig interpretiert werden, geben aber durchaus Anlass, inhaltliche Konsequenzen für den praktischen Umgang mit Bewegungsbeurteilungsprozessen abzuleiten oder sogar neue Forschungsfragen und Hypothesen zu stellen, die vor dieser Studie in dieser Form und Detailliertheit noch nicht möglich gewesen wären.

8 Appendix

Dunham Model

Body orientation: _____

Preparation: _____

Feet: _____

Knees: _____

Hips: _____

Trunk: _____

Shoulders: _____

Arms: _____

Hands: _____

Head: _____

Execution: _____

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____

Abb. 113: Task sheet für die Bewegungsbeobachtung des Überkopfwurfes nach Dunham (1994)

Tab. 44: Fragebogen zur Prüfung des Fachwissens über die leichtathletische Schlagwurfbewegung

Fachwissen Ballwurf, Seite 1 von 5

Liebe Probandin, lieber Proband,
diese Prüfung wird maschinell ausgewertet. Markieren Sie bitte bei den Multiple-Choice-Fragen Ihre Antwort(en) in der folgenden Weise: ○ ⊗ ○ .
Wenn Sie eine Antwort korrigieren möchten, füllen Sie bitte die falsch markierte Antwort vollständig aus, ungefähr so: ○ ● ⊗ . Sollten Sie ein ausgefülltes Feld doch wieder wählen wollen, kreisen Sie dieses ein.
Die Fünfecke unter den Textfragen werden nur vom Korrektor ausgefüllt.
Vielen Dank für Ihre Teilnahme an der Untersuchung!



Matrikelnummer

Bitte schreiben Sie unter diesen Text Ihren Namen und kreuzen Sie im rechten Bereich Ihre Matrikelnummer an (erste Ziffer = oberste Zeile).		0	1	2	3	4	5	6	7	8	9
	_____	<input type="radio"/>									
	_____	<input type="radio"/>									
	_____	<input type="radio"/>									
	_____	<input type="radio"/>									
	_____	<input type="radio"/>									

Fachwissen Ballwurf

Anlauf

1.1 Kreuzen Sie die Aussagen an, welche bezüglich der **Ballrückführung** richtig sind. (4 Punkte)

- Die Ballrückführung erfolgt im letzten Anlaufschritt.
- Der Gegenarm wird mit der Ballrückführung am Körper angelegt und zeigt nach unten zum Boden.
- Während der Ballrückführung dreht der Athlet die Schulterachse von einer frontalen in eine parallele Position zur Wurfrichtung.
- Der Wurfarm wird auf Schulterhöhe zurückgeführt und dann im Ellbogengelenk gestreckt.

Impulsschritt

2.1 Kreuzen Sie die Aussagen an, welche bezüglich des **Impulsschrittes** richtig sind. (4 Punkte)

- Der Impulsschritt wird durch ein ziehendes Aufsetzen und einen kräftigen Abdruck im drittletzten Bodenkontakt eingeleitet.
- Der Impulsschritt wird mit einer hohen Flugkurve weit nach vorne ausgeführt.
- Während des Abdrucks zum Impulsschritt schwingt das Bein der Wurfarmseite am anderen Bein vorbei nach vorne.
- Der Oberkörper wird durch den Impulsschritt in eine leichte Vorlage gebracht.

Wurfauslage

3.1 Kreuzen Sie die Aussagen an, welche bezüglich der **Druckbeinbewegung** in der Wurfauslage richtig sind. Beachte: Druckbein = hinteres Bein in der Abwurfposition, Stemmbein = vorderes Bein (4 Punkte)

- Das Druckbein wird mit einer Beugung im Kniegelenk aufgesetzt.
- Die Fußspitze des Druckbeins sollte beim Aufsetzen (entgegen der Wurfrichtung) nach hinten zeigen.
- Nach dem Aufsetzen des Druckbeins wird zuerst die Hüfte nach vorne gedreht und beschleunigt.
- Nach dem Aufsetzen drückt das Druckbein in Wurfrichtung nach vorne gegen das Stemmbein.

3.2 Kreuzen Sie die Aussagen an, welche bezüglich des **Aufsetzens des Stemmbeins** richtig sind. (4 Punkte)

- Die Fußspitze des Stemmbeins zeigt quer zur Wurfrichtung.
- Das Stemmbein wird auf dem Fußballen aufgesetzt.
- Das Stemmbein wird gestreckt aufgesetzt.
- Das Stemmbein sollte während des Abwurfs möglichst wenig gebeugt werden.

3.3 Entscheiden Sie, welche Aussagen bezüglich der **räumlich-zeitlichen Ausführung des Stemmschrittes** richtig sind. (4 Punkte)

- Noch bevor das Druckbein aufsetzt, wird das Stemmbein wieder in Richtung Druckbein beschleunigt.
- Das Stemmbein wird aufgesetzt, nachdem sich das Druckbein im Abwurf vom Boden löst.
- Der zeitliche Abstand zwischen dem Setzen des Druckbeins und dem Setzen des Stemmbeins sollte möglichst groß sein.
- Der Stemmschritt ist (räumlich betrachtet) länger als der Impulsschritt.



Abwurfbewegung

<p>4.1 Kreuzen Sie an, welche Aussagen bezüglich der Körperverwindung im Abwurf richtig sind. (4 Punkte)</p> <p><input type="checkbox"/> Beim Aufsetzen des Stemmbeins wird die Wurfhand mit möglichst gestrecktem Ellbogen zurückgehalten.</p> <p><input type="checkbox"/> Der Schultergürtel befindet sich zu Beginn der Abwurfbewegung parallel zur Wurfrichtung.</p> <p><input type="checkbox"/> Die Körperverwindung im Abwurf startet mit dem Vorbringen der wurfseitigen Hüfte.</p> <p><input type="checkbox"/> Die Bogenspannung wird durch aktives Aufdrehen der Wurf Schulter nach hinten erreicht.</p>
<p>4.2 Kreuzen Sie an, welche Aussagen bezüglich der Ausführung der Wurfbewegung richtig sind. (4 Punkte)</p> <p><input type="checkbox"/> Der Oberkörper knickt im Abwurf seitlich zur Wurfarmgegenseite ab.</p> <p><input type="checkbox"/> Zu Beginn der Wurfarmbeschleunigung eilt der Ellbogen der Wurfhand voraus.</p> <p><input type="checkbox"/> Der Gegenarm bleibt beim Abwurf vor dem Körper.</p> <p><input type="checkbox"/> Die Wurfhand befindet sich kurz vor Verlassen des Balles über Kopfhöhe.</p>

Abfangen

<p>5.1 Kreuzen Sie die Aussagen an, welche bezüglich des Abfangens der Wurfbewegung richtig sind. (4 Punkte)</p> <p><input type="checkbox"/> Im Moment des Verlassens des Balls von der Wurfhand befindet sich der Oberkörper in Rücklage.</p> <p><input type="checkbox"/> Im Moment des Verlassens des Balls von der Wurfhand befindet sich die Schulterachse quer zur Wurfrichtung.</p> <p><input type="checkbox"/> Das Stemmbein bleibt nach dem Abfangen möglichst gestreckt.</p> <p><input type="checkbox"/> Das Druckbein hat im Moment des Ballabwurfs keinen Kontakt mehr mit dem Boden.</p>
--

Biomechanik

<p>6.1 Kreuzen Sie die Aussagen an, welche bezüglich der letzten drei Anlaufschritte richtig sind. (4 Punkte)</p> <p><input type="checkbox"/> Mit dem Impulsschritt beginnt das Abbremsen der Horizontalgeschwindigkeit.</p> <p><input type="checkbox"/> Das Druckbein wird im vorletzten Bodenkontakt ungefähr auf die vertikale Projektion des Körperschwerpunktes auf dem Boden aufgesetzt.</p> <p><input type="checkbox"/> Je weiter der Stemmbeinfuß im Moment des Aufsetzens des Druckbeins vor diesem ist, desto schneller kann das Stemmbein in der Folge zum Boden gesetzt werden.</p> <p><input type="checkbox"/> Der Gegenarm unterstützt durch sein langes Verharren in Wurfrichtung das späte Drehen der Schulterachse in die Frontalposition.</p>
<p>6.2 Kreuzen Sie die Aussagen an, welche bezüglich der Bewegung des Wurfarms richtig sind. (4 Punkte)</p> <p><input type="checkbox"/> Am Ende der Wurfbewegung wird der Ellbogen kurz vor seiner Streckung abgebremst.</p> <p><input type="checkbox"/> Die Handgelenkstreckung übt den letzten Beschleunigungsimpuls auf das Wurfgerät aus.</p> <p><input type="checkbox"/> Aus der Wurfauslage erfolgt das Beschleunigen und Abbremsen der einzelnen Wurfarmgelenke zeitlich betrachtet von distal (fern vom Körperzentrum) nach proximal (nah am Körperzentrum).</p> <p><input type="checkbox"/> Der Ellbogen des Wurfarms wird beim Abwurf zeitgleich mit der Wurfhand nach vorne gebracht.</p>
<p>6.3 Kreuzen Sie die Aussagen an, welche bezüglich der Körperverwindung und Bogenspannung richtig sind. (4 Punkte)</p> <p><input type="checkbox"/> Kurz vor dem Aufsetzen des Stemmbeins erreicht die wurfseitige Hüfte ihr Geschwindigkeitsmaximum.</p> <p><input type="checkbox"/> Die entgegengesetzt wirkende Druck- und Stemmbeinarbeit führt zur Verwindung zwischen Hüft- und Schulterachse.</p> <p><input type="checkbox"/> Mit dem Vorbringen der Wurfarmschulter bei gleichzeitiger Verzögerung des Wurfarms ist der Aufbau der Bogenspannung abgeschlossen.</p> <p><input type="checkbox"/> Der Wurfarm sollte solange verzögert möglichst gestreckt hinter dem Körper bleiben, bis die wurfarmseitige Schulter nach der Beschleunigung wieder abgebremst wird.</p>
<p>6.4 Kreuzen Sie die Aussagen an, welche bezüglich des Abfangens der Wurfbewegung richtig sind. (4 Punkte)</p> <p><input type="checkbox"/> Im Moment des Abwurfs sollte der Körperschwerpunkt den Aufsatzpunkt des Stemmbeins bereits überschritten haben.</p> <p><input type="checkbox"/> Die Rückführung des Gegenarms an den Körper beim Abwurf unterstützt die Längsachsenrotation des Oberkörpers.</p> <p><input type="checkbox"/> Mit einem Schritt über das Stemmbein hinweg kann die noch vorhandene Horizontalgeschwindigkeit aus dem Abwurf abgefangen werden.</p> <p><input type="checkbox"/> Das Stemmbein bleibt beim Abfangen möglichst lang hinter dem Körper zurück, um den Körperschwerpunkt möglichst weit hinten zu halten.</p>



Fehlerwissen

Im folgenden Aufgabenteil wird nach Fehlern gefragt, **welche bei Anfängern häufig auftauchen** können. Achten Sie vor allem darauf, **die schwerwiegendsten Fehler** zu nennen bzw. solche, die Ihrer Meinung nach **bei Anfängern sofort korrigiert** werden müssten.

Konzentrieren Sie sich dabei darauf, nicht ausschließlich die Auswirkung eines Fehlers zu nennen (z.B. schwacher Abwurf), sondern die dafür verantwortliche **Ursache** zu beschreiben. Eine stichpunktartige Beantwortung reicht!

7.1 Nennen Sie **fünf** typische Fehler, die den **Anlauf zum Schlagwurf** betreffen und bei Anfängern häufig vorkommen. Die Wurfauslage gehört **nicht** mehr zur Beantwortung dieser Frage! (10 Punkte)

Erreichte Punktzahl:

0
 1
 2
 3
 4
 5
 6
 7
 8
 9
 10

7.2 Nennen Sie **fünf** typische Fehler, die den **Abwurfvorgang** (beginnend mit dem Aufsetzen des Stemmbeins!) betreffen und bei Anfängern häufig vorkommen. (10 Punkte)

Erreichte Punktzahl:

0
 1
 2
 3
 4
 5
 6
 7
 8
 9
 10

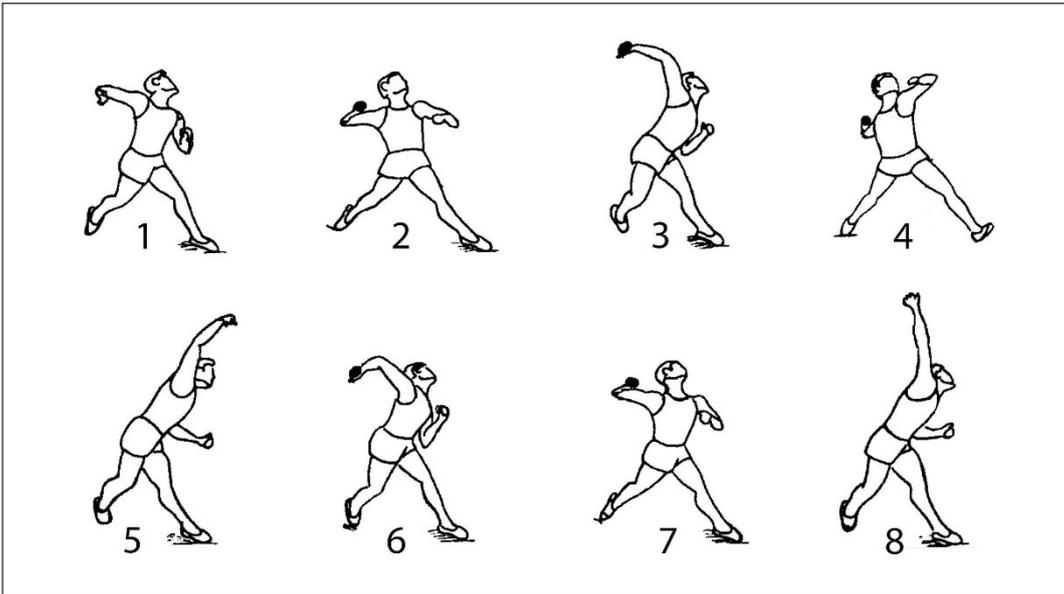


7.3 Nennen Sie **zwei** zentrale Fehler, die das **Abfangen** des Schlagwurfs (beginnend mit dem Verlassen des Balles aus der Wurfhand) betreffen und bei Anfängern häufig vorkommen. (4 Punkte)

Erreichte Punktzahl:

0 1 2 3 4

Bildreihe



8.1 Ordnen Sie die obigen Bilder nach ihrer chronologischen Reihenfolge im Bewegungsverlauf. Notieren Sie die entsprechenden Zahlen in der richtigen Reihenfolge (von links nach rechts). (8 Punkte)

Erreichte Punktzahl:

0 1 2 3 4 5 6 7 8

Tab. 45: Synopse der Fehlerbilder für die Phasen Anlauf (Frage 7.1) und Abwurf (Frage 7.2)

	Frage 7.1: Anlauf zum Schlagwurf (Wurfauslage gehört nicht mehr zur Frage!)
1. Fehlerbilder Anlauf zyklisch	
1x	Anlaufgeschwindigkeit zu gering / ungenügende Geschwindigkeitssteigerung
2x	Anlaufgeschwindigkeit zu hoch / Geschwindigkeitsabfall mit Beginn oder während des zyklischen Anlaufteils
3	Unruhige Arm-/Speerführung bzw. betontes Vor- und Rückführen des Wurfarmes im zyklischen Anlaufteil
4x	Speer nicht hoch genug, hängendes Speerende >> 3.5.
5	Feste, starre Oberkörper- und Armhaltung
2. Fehlerbilder Anlauf azyklisch	
1x	Ungenügende Körperrücklage
2x	Zu geringe Körper-/Oberkörperrücklage nach Impulsschritt (bzw. beim Aufsetzen des Druckbeins)
3x	Körpervorlage
4x	Zu große Oberkörperrücklage beim Aufsetzen des Druckbeins
5x	Keine Achsenparallelität von Schulter-, Hüft- und Speerachse beim Setzen des Druckbeins
6x	Vorzeitiges Anziehen bzw. Einbeugen des Wurfarmes im Anlauf / nach der Rückführung
7x	Zu hoher, sprunghafter Impulsschritt (einhergehend mit Horizontalgeschwindigkeitsverlusten beim Abdruck zum Impulsschritt und Aufsetzen des Druckbeins)
8x	Impulsschritt zu kurz / ohne Doppelwechsel der Beine
9x	Falscher Abwurfrhythmus
10x	Impulsschritt führt nicht in die Wurfauslage, das Gewicht liegt zwischen den Beinen und nicht auf dem rechten Bein, das linke Bein kann nicht als Stemmbein eingesetzt werden
11x	Schlechte Schrittgestaltung
12x	Abfall der Geschwindigkeit zu Beginn des Fünf-Schritt-Rhythmus bzw. mangelnde oder fehlende Steigerung der Geschwindigkeit
13x	Impulshopser/Nachstellschritt, Nachstellhupf
14x	Koordinative Probleme beim seitlichen Laufen mit zurückgenommenem Speer
15x	Füße zeigen nicht in Laufrichtung / Aufsetzen des rechten Fußes quer zur Wurfrichtung

3. Fehlerbilder Rückführung Wurfgerät	
1	Sinkende Geschwindigkeit vor der Speerabnahme
2x	Zu langsame oder zu späte Speerrücknahme
3x	Unkoordinierte Speerrückführung
4x	Unvollständige bzw. ungenügende Speerrückführung
5x	Falsche Wurfarm- und Speerhaltung nach der Rücknahme

Frage 7.2: Abwurfbewegung	
1. Gesamtkoordination	
1x	Beim Abwurf falsches Bein vorne
2. Oberkörperposition beim Druckbeinaufsatz	
1x	Zu geringe Oberkörperrücklage
2x	Zu große Rücklage
3x	Schulterachse zeigt beim Setzen des Druckbeins nicht in Wurfrichtung
3. Druckbeinaufsatz	
1x	Zu inaktives (vorwiegend nachgebendes) Druckbein / passive Druckbeinarbeit
2	Fehlende Dreh-Stoßbewegung des Druckbeins
4. Stemmbeinaufsatz	
1	„Fallen“ auf das linke Bein
2x	Gebeugtes Aufsetzen des Stemmbeins bzw. Nachgeben im Kniegelenk / „sitzende“ Abwurfhaltung
3x	Flugphase im Stemmschritt
4	Setzen des Stemmbeins von oben
5	Zu kurzer Stemmschritt
6	Zu langer Stemmschritt
7x	Zu schwache Stemmbeinarbeit (zu geringe Stemmwirkung) / ungenügende Hebelwirkung des Stemmbeins

5. Kopplung der Teilbewegungen	
1x	Keine Auflösung der Wurfauslage nach dem Prinzip der Koordination der Teilimpulse
2x	Ungenügende Wurfverzögerung (mangelndes Zurückhalten der Wurfhand) beim Setzen des Stemmbeins / Wurfarm bzw. -schulter werden vorzeitig angezogen
3	Ungenügende Eindrehbewegung des rechten Beines zu Beginn der Abwurfbewegung
4	Mangelndes Blockieren der linken Körperseite
5	„Standabwurf“
6	Flüchtiger Abwurf
7. Bogenspannung	
1x	Mangelhafte, zu geringe oder fehlende Bogenspannung
2x	Keine ausreichende Verwringung zwischen Schulter- und Hüftachse beim Setzen des Stemmbeins
8. Wurfarmbewegung	
1	Falsche Hand- und/oder Speerposition
2x	Gebeugter Wurfarm / ungenügende Streckung im Wurfarmellenbogen beim Setzen des Druckbeins
3x	Wurfhand befindet sich zu Beginn der Abwurfbewegung (deutlich) unter Schulterhöhe
4x	Tiefe, seitliche Armführung / zu tiefer Weg der Wurfhand
5x	Wurfarm wird seitlich neben der Schulter vorbeigezogen (Schlenzwurf)
6x	Mangelhafte Unterarmschleuder
7	Schlechte Speerführung
9. Körperposition beim Abwurf	
1	Angezogenes Druckbein beim Abwurf
2x	Zu starke seitliche Ausweichbewegungen
3	Starkes Abknicken des Oberkörpers zur Wurfarmgegenseite im Abwurf
4x	Fehlende Körperstreckung beim Abwurf / Ausweichen der Hüfte nach hinten
5	Hüfte nicht fixiert, Ausgleichbewegung entgegen der Wurfrichtung
10. Abflugwinkel	
1	Zu steiler Abflugwinkel des Geräts
2	Zu flacher Abwurfwinkel

Tab. 46: Erwartungshorizont für das Fehlerwissen der Phase Anlauf azyklisch

1. Fehlerbilder Anlauf azyklisch	Autor
A) Oberkörperposition/Körperlage	
1. Ungenügende Körperrücklage	Bauersfeld
2. Zu geringe Körper-/Oberkörperrücklage nach Impulsschritt (bzw. beim Aufsetzen des Druckbeins) <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ zu kurzer Impulsschritt >> lässt keine Zeit für die Beine, den Oberkörper zu überholen ▪ zu passiver Impulsschritt ▪ die Hüfte überholt Wurf Schulter nicht ▪ kein Überkreuzen der Beine während des Impulsschritts ▪ zu frühe Abwurfaktivität (Vorbringen der Wurf Schulter) ▪ angestrebter Abwurfwinkel zu gering ▪ Blickorientierung zu sehr nach unten (auf Landepunkt) ▪ Werfer ist nicht hinreichend um einen langen Beschleunigungsweg nach dem Setzen des Druckbeins bemüht 	Weiler, Bauersfeld, Kunz, Strüder
3. Körpervorlage <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ Beine laufen dem Rumpf nicht voraus ▪ mangelhafte Hüft- und Schulterbeweglichkeit ▪ Rhythmusfehler 	Stahl
4. Zu große Oberkörperrücklage beim Aufsetzen des Druckbeins <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ zu hoher (bzw. großer) Impulsschritt ▪ angestrebter Abwurfwinkel zu groß ▪ Blickorientierung zu sehr nach oben ▪ Werfer ist nicht hinreichend um ein schnelles Weiterarbeiten auf dem Druckbein bemüht 	Strüder
5. Keine Achsenparallelität von Schulter-, Hüft- und Speerachse beim Setzen des Druckbeins <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ Abknicken im Oberkörper ▪ ungenügendes Beschleunigen der Beine im Impulsschritt 	Killing 2011
B) Wurfarm	
6. Vorzeitiges Anziehen bzw. Einbeugen des Wurfarmes im Anlauf / nach der Rückführung <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ mangelnde Beweglichkeit ▪ unzureichende Kraftfähigkeiten 	Bauersfeld, Killing 2011

C) Impulsschritt	
7. Zu hoher, sprunghafter Impulsschritt (einhergehend mit Horizontalgeschwindigkeitsverlusten beim Abdruck zum Impulsschritt und Aufsetzen des Druckbeins) <u>>> Mögliche Ursachen:</u>	Strüder, Bauersfeld, Killing 2011
<ul style="list-style-type: none"> ▪ zu hohe Anlaufgeschwindigkeit ▪ Anfänger führt bewusst einen Sprung aus ▪ Abdruck nach oben statt nach vorn ▪ Schwungbein (das in der Folge zum Druckbein wird) wird im Abdruck zum Impulsschritt zu hoch geführt ▪ Werfer will zu großen Impulsschritt ausführen 	
8. Impulsschritt zu kurz / ohne Doppelwechsel der Beine <u>>> Mögliche Ursachen:</u>	Bauersfeld, Killing 2011
<ul style="list-style-type: none"> ▪ unzureichende Bein-/Rumpfkraft ▪ Oberkörpervorlage ▪ zu hohe Eingangsgeschwindigkeit 	
9. Falscher Abwurfrythmus <u>>> Mögliche Ursachen:</u>	Kunz
<ul style="list-style-type: none"> ▪ zu kurzer Impulsschritt ▪ zu lange Schrittauslage ▪ verzögertes Aufsetzen des linken Beins 	
10. Impulsschritt führt nicht in die Wurfauslage, das Gewicht liegt zwischen den Beinen und nicht auf dem rechten Bein, das linke Bein kann nicht als Stemmbein eingesetzt werden	DSHS
11. Schlechte Schrittgestaltung	Kunz
D) Geschwindigkeitsabfall	
12. Abfall der Geschwindigkeit zu Beginn des Fünf-Schritt-Rhythmus bzw. mangelnde oder fehlende Steigerung der Geschwindigkeit <u>>> Mögliche Ursachen:</u>	Weiler, DSHS, Bauersfeld, Strüder, Killing 2011
<ul style="list-style-type: none"> ▪ zu hohe Anlaufgeschwindigkeit / zyklischer Anlauf zu schnell ▪ zu hoher Auftaktschritt / zu hoher Übergangshopser zwischen zyklischem und azyklischem Anlaufteil ▪ mangelnde Koordination des zyklischen mit dem azyklischen Bewegungsabschnitt ▪ zu langsame oder zu späte Speerrücknahme ▪ koordinative Probleme beim seitlichen Laufen mit zurückgenommenem Speer ▪ zu starker Abdruck mit dem linken Bein ▪ mangelnde Bewegungskoordination der Arme und Beine 	
E) Impulshopser	
13. Impulshopser/Nachstellschritt, Nachstellhupf	Haberkorn, Bauersfeld, Stahl

F) Lauf- und Fußkoordination	
14. koordinative Probleme beim seitlichen Laufen mit zurückgenommenem Speer (z. B. Schulterachse in Laufrichtung, Füße quer zur Laufrichtung, eher Seitgalopp statt Kreuzschritt, ungünstige Blickrichtung o.Ä.)	Strüder
15. Füße zeigen nicht in Laufrichtung / Aufsetzen des rechten Fußes quer zur Wurfrichtung <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ fehlender Schwungbeineinsatz beim Impulsschritt ▪ kein Vordrehen des rechten Fußes vor der Landung ▪ zu hohe Konzentration auf den Hüfteinsatz ▪ Beweglichkeitsdefizite 	Killing 2011, Kunz

Tab. 47: Erwartungshorizont für das Fehlerwissen der Phase Wurfgerätrückführung

2. Fehlerbilder Rückführung	<i>Autor</i>
A) Geschwindigkeitsverlust im Anlauf	
1. Sinkende Geschwindigkeit vor der Speerabnahme <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ zu hohe Anfangsgeschwindigkeit ▪ kein Ballenlauf 	Killing, 2011
B) Fehlerhafte Speerrückführung	
2. Zu langsame oder zu späte Speerrücknahme <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ mangelnder Schrittrhythmus ▪ falsche Bewegungsvorstellung hinsichtlich des zeitlichen Ablaufs der Rückführung im 5er-Rhythmus ▪ koordinative Probleme beim seitlichen Laufen mit zurückgenommenem Speer 	Strüder
3. Unkoordinierte Speerrückführung <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ bogenförmige Rückführung des Wurfarmes ▪ fehlende Rumpfdrehung ▪ seitliches Ausweichen des Armes 	Bauersfeld, Kunz, Stähl
4. Unvollständige bzw. ungenügende Speerrückführung <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ mangelhafte Beweglichkeit im Hüft- und Schulterbereich ▪ zu späte Rückführung des Speers ▪ keine aktive Streckung des Wurfarmes ▪ Abknicken im Handgelenk ▪ Gegenarm nicht vorgebracht ▪ unzureichende Bewegungsvorstellung ▪ Beweglichkeitsdefizite 	Killing 2011, Weiler, Stähl, Hab- erkorn, Bau- ersfeld

<p>5. falsche Wurfarm- und Speerhaltung nach der Rücknahme: gebeugter Arm, Wurfarm tief (Spitze zu hoch), Handgelenk aufgeklappt (Spitze zu hoch und meist zu seitlich vom Kopf), Speerspitze zu tief (Brust-, Schulter-, Halshöhe)</p> <p>>> <u>Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ Oberkörper/Schulter kippt ab ▪ Arm wird zu tief geführt 	Haberkorn, Bauersfeld, DSHS, Killing 2011
---	---

Tab. 48: Erwartungshorizont für das Fehlerwissen der Phase Abwurf

3. Fehlerbilder Abwurf	<i>Autor</i>
A) Gesamtkoordination	
1. Beim Abwurf falsches Bein vorne	Haberkorn
B) Oberkörperposition beim Druckbeinaufsatz	
2. Mangelhafte Wurfauslage: zu geringe Oberkörperrücklage	Haberkorn
3. Zu große Rücklage bei Setzen des Druckbeins >> <u>Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ zu passives Druckbein vor dem Aufsetzen ▪ falsche Bewegungsvorstellung 	Killing 2011
4. Schulterachse zeigt beim Setzen des Druckbeins nicht in Wurfrichtung >> <u>Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ zu langsame oder zu späte Speerrücknahme ▪ Werfer ist nicht hinreichend um einen langen Beschleunigungsweg in der Abwurfbewegung bemüht (kein aktives Zurückführen der Wurfhand) 	Strüder, Stähl
C) Druckbeinaufsatz	
5. Zu inaktives (vorwiegend nachgebendes Druckbein) / passive Druckbeinarbeit >> <u>Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ mangelnde Reaktivität im Druckbein ▪ Aufsetzen des Druckbeins quer zur Laufrichtung (mehr als 45° zur Laufrichtung) ▪ zu gestrecktes Aufsetzen des Druckbeins ▪ Druckbein kann im Kniegelenk nicht ausreichend fixiert werden ▪ Anlaufgeschwindigkeit zu hoch 	Strüder, Killing 2011
6. Fehlende Dreh-Stoßbewegung des Druckbeins >> <u>Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ passives Fußaufsetzen quer zur Wurfrichtung ▪ kein Weiterdrehen des rechten Fußes, mangelnde Reaktivkraft 	Kunz

D) Stemmbeinaufsatz	
7. ‚Fallen‘ auf das linke Bein	Bauersfeld
8. Gebeugtes Aufsetzen des Stemmbeins bzw. Nachgeben im Kniegelenk / „Sitzende“ Abwurfhaltung <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ Oberkörper bereits zu weit vorne ▪ ungenügend entwickelte Kraftfähigkeiten ▪ Stemmbein von oben gesetzt ▪ zu kurzer Stemmschritt ▪ zu geringe Körperrücklage ▪ zu passives Aufsetzen des Stemmbeins ▪ ungenügender oder fehlender Stemmbeineinsatz (Hebelwirkung!) 	Killing 2011, Weiler, Haberkorn
9. Flugphase im Stemmschritt (hauptsächlich bei Anfängern) <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ zu hohe Stemmbeinführung ▪ zu langer Stemmschritt ▪ mangelndes Bemühen um schnelles Weiterarbeiten des Druckbeins gegen das Stemmbein 	Strüder
10. Setzen des Stemmbeins von oben <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ ungenügende Beugewinkel im Druckbein (insgesamt zu passives Druckbein) ▪ zu kurzer Stemmschritt 	Killing 2011
11. Zu kurzer Stemmschritt <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ unzureichendes Überholen im Impulsschritt 	Killing 2011, Stähl
12. Zu langer Stemmschritt <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ zu kurzer Impulsschritt ▪ unzureichende Fuß-/Beinaktivität ▪ eventuell zu hohe Anlaufgeschwindigkeit 	Killing 2011
13. Zu schwache Stemmbeinarbeit (zu geringe Stemmwirkung) / ungenügende Hebelwirkung des linken Beins <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ zu großer Stemmwinkel (zwischen Bein und Horizontaler) durch zu kurzen Stemmschritt ▪ zu starkes Nachgeben im Kniegelenk des Stemmbeins durch: zu geringen Stemmwinkel durch zu langen Stemmschritt, Aufsetzen des Stemmbeins auf dem Fußballen oder mit gebeugtem Knie ▪ mangelnde Streckkraft im Stemmbein ▪ zu hohe Anlaufgeschwindigkeit ▪ zu geringe Oberkörperücklage beim Setzen des Druckbeins 	Strüder, Bauersfeld
E) Kopplung der Teilbewegungen	
14. Keine Auflösung der Wurfauslage nach dem Prinzip der Koordination der Teilimpulse <u>>> Mögliche Ursachen:</u> <ul style="list-style-type: none"> ▪ alle Segmente der Wurfarmseite bewegen sich gleichzeitig nach vorne ▪ vorzeitiger Einsatz des Oberkörpers und Wurfarmes 	Bauersfeld, DSHS

<p>15. Ungenügende Wurfverzögerung (mangelndes Zurückhalten der Wurfhand) beim Setzen des Stemmbeins / Wurfarm bzw. -schulter werden vorzeitig angezogen <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ zu lange Dauer der Phase ‚Setzen Druckbein bis Setzen Stemmbein‘ ▪ ungenügende Entspannungsfähigkeit von Rumpf/Schultergürtel ▪ mangelnde Flexibilität im Schulterbereich ▪ ungenügende Streckung im Wurfarmellenbogen beim Setzen des Druckbeins ▪ kein Eindrehen des rechten Beines ▪ zu früher Armeinsatz ▪ falsche Bewegungsvorstellung 	<p>Strüder, Killing 2011, Kunz</p>
<p>16. Ungenügende Eindrehbewegung des rechten Beines zu Beginn der Abwurfbewegung</p>	<p>Bauersfeld</p>
<p>17. Mangelndes Blockieren der linken Körperseite <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ passives Aufsetzen/Beugen des linken Beines ▪ fehlende Körperspannung ▪ Zurückreißen des linken Armes 	<p>Kunz</p>
<p>18. ‚Standabwurf‘ <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ zu hohe, nicht umsetzbare Anlaufgeschwindigkeit 	<p>Haberkorn</p>
<p>19. Flüchtiger Abwurf <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ Anlaufgeschwindigkeit beim Abwurf nicht umsetzbar ▪ kein Aufbau der Bogenspannung ▪ mangelhafte Ausführung des Impulsschrittes (-hopsers), evtl. Nachstellschritt 	<p>Haberkorn</p>
F) Bogenspannung	
<p>20. Mangelhafte, zu geringe oder fehlende Bogenspannung <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ zu kurzer Impulsschritt ▪ unzureichende Rumpfrücklage ▪ „Armwurf“ (zu früher Wurfarmeinsatz) / mangelndes Zurückhalten der Wurfhand ▪ kein sofortiges Eindrehen des hinteren Beines ▪ mangelhafte Hüft- und Schulterbeweglichkeit ▪ passive Druckbeinarbeit ▪ zu schwache Stemmbeinarbeit 	<p>Bauersfeld, Haberkorn, Stähl, Strüder</p>
<p>21. Keine ausreichende Verwringung zwischen Schulter- und Hüftachse beim Setzen des Stemmbeins <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ zu frühzeitige Aktivierung des Wurfarms und/oder der Wurfarmschulter ▪ zu geringe Aktivität des Druckbeins zum Setzen des Stemmbeins 	<p>Killing 2011</p>
G) Wurfarmbewegung	
<p>22. Falsche Hand- und/oder Speerposition (Spitze weicht vom Kopf ab, Ende berührt Boden) <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ zu kurzer Impulsschritt ▪ häufig fehlende Wurfarm- und Speerfixierung 	<p>Haberkorn</p>

<p>23. Gebeugter Wurfarm / ungenügende Streckung im Wurfarmellenbogen beim Setzen des Druckbeins <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ Gegenarm wurde zu wenig vorgebracht („Schließen“) ▪ Vorwegnehmen der Abwurfbewegung: Werfer beginnt Schulterachsenrotation („Öffnen“) oder Wurfarmeinsatz zu früh (vor Erreichen der Wurfauslage) ▪ koordinative Probleme beim seitlichen Laufen mit zurückgenommenem Speer ▪ keine Wurfarmstreckung bei Ballrücknahme 	<p>Strüder, Haberkorn</p>
<p>24. Wurfhand befindet sich zu Beginn der Abwurfbewegung (deutlich) unter Schulterhöhe <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ zu tiefes Führen der Wurfhand nach der Speerrücknahme ▪ zu große Oberkörperrücklage beim Aufsetzen des Druckbeins 	<p>Strüder</p>
<p>25. Tiefe, seitliche Armführung / Zu tiefer Weg der Wurfhand (führt häufig zu einem stark bogenförmigen Beschleunigungsweg außen herum) <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ kein Vor-hoch-Bringen des Ellenbogens und Eindrehen des Wurfarms (Ellenbogen wird während der Abwurfbewegung zum Teil unter Schulterhöhe geführt) ▪ tiefer, angezogener Wurfarm ▪ fehlende Wurfarmverzögerung ▪ schlechte Schulterbeweglichkeit 	<p>Strüder, Kunz</p>
<p>26. Wurfarm wird seitlich neben der Schulter vorbeigezogen (Schlenzwurf) <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ Ellbogen zieht zu wenig nach vorn-oben >> Vorbringen des Ellbogens in Hals- oder Schulterhöhe ▪ ungenügende Stemmwirkung ▪ mangelhafte Schulterbeweglichkeit 	<p>Stahl, Haberkorn</p>
<p>27. mangelhafte Unterarmschleuder (Mangelhafte Endbeschleunigung des Speers) <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ zu tiefer Weg der Wurfhand ▪ zu frühes Loslassen ▪ mangelnde Reaktivkraft im Ellbogenstrecker ▪ zu geringe anfängliche Ellbogenstreckung durch vorzeitiges und zu starkes Beugen des Ellbogengelenks und durch eine mangelhafte Eindrehbewegung, ▪ mangelnde Schnellkraftfähigkeiten 	<p>Strüder, Killing 2011</p>
<p>28. Schlechte Speerführung <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ angezogener Arm ▪ abgekipptes Handgelenk ▪ Speer quer zur Wurfrichtung 	<p>Kunz</p>
<p>H) Körperposition beim Abwurf</p>	
<p>29. Angezogenes rechtes Bein beim Abwurf <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ zu frühes Lösen des rechten Fußes vom Boden und Vorziehen (kein Beinstütz) 	<p>Haberkorn</p>
<p>30. Zu starke seitliche Ausweichbewegungen (aus der Rückenansicht zu erkennen) (z. B. Laufweg der letzten Schritte, Vor- und Rückenneigung des Rumpfes während der Abwurfvorbereitung oder des Abwurfs) <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ Werfer ist nicht hinreichend um eine geradlinige Abwurfbewegung bemüht 	<p>Strüder</p>

<p>31. Starkes Abknicken des Oberkörpers zur Wurfarmgegenseite im Abwurf <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ fehlender Impuls aus dem rechten Bein ▪ mangelnde Körperspannung ▪ früher Armeinsatz ▪ Blick zum Boden ▪ fehlende Rumpfstreckung und -fixierung im linken Hüftgelenk, meist verbunden mit mangelhafter Stemmbeinstreckung 	<p>Bauersfeld, Kunz, Haberkorn</p>
<p>32. Fehlende Körperstreckung beim Abwurf / Ausweichen der Hüfte nach hinten <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ ‚eingeknicktes‘ Stemmbein ▪ frühzeitiges Lösen des rechten Fußes vom Boden und Vorziehen ▪ mangelhafte Rumpfrücklage zu Beginn der Wurfauslage ▪ zu großer Stemmschritt ▪ zu wenig Kraft in den Beinen und im Rumpf ▪ mangelhafte Hüft- und Schulterbeweglichkeit 	<p>Haberkorn, Stähl</p>
<p>33. Hüfte nicht fixiert, Ausgleichbewegung entgegen der Wurfrichtung <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ ungenügendes Eindrehen durch das Druckbein 	<p>Killing 2011</p>
<p>1) Abflugwinkel</p>	
<p>34. Zu steiler Abflugwinkel des Gerätes (über 40 Grad) <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ mangelnde Beschleunigung des Oberkörpers, des Ellenbogens und der Hand nach vorn ▪ zu große Oberkörperrücklage beim Setzen des Druckbeins ▪ Wurfhand befindet sich zu Beginn der Abwurfbewegung unter Schulterhöhe ▪ zu geringe Anlaufgeschwindigkeit ▪ zu starke Streckbewegung aus dem Stemmbein beim Abwurf ▪ zu großer Anstellwinkel ▪ zu früher und betonter Armeinsatz ▪ zu langer Stemmschritt ▪ Sprungabwurf 	<p>Bauersfeld, Strüder, Killing 2011, Kunz</p>
<p>35. Zu flacher Abwurfwinkel (unter 32 Grad) <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ zu geringer Stemmbeineinsatz ▪ zu kurzer Stemmschritt ▪ zu früher Einsatz von Wurfarm/-schulter beim Setzen des Stemmbeins ▪ übermäßiges Bestreben um eine hohe Abfluggeschwindigkeit zu Ungunsten des Abflugwinkels (übermäßige Rumpfbeugung nach vorn) ▪ zu geringe Oberkörperrücklage beim Setzen des Druckbeins ▪ zu tiefer Weg der Hand und dessen Ursachen ▪ Hüftknick beim Abwurf ▪ Einknicken des linken Beins 	<p>Strüder, Killing 2011, Kunz</p>
<p>36. Speer zu stark herausgehoben (Anstellwinkel deutlich kleiner als Abwurfwinkel) <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ zu vertikale Bewegungsrichtung des Oberkörpers und Wurfarms bei optimalem Anstellwinkel 	<p>Killing 2011</p>

<p>37. Speer zu stark angestellt (Anstellwinkel um mehr als 5 Grad größer als der Abwurfwinkel) / Speer steht in der Luft <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ Abknicken im Handgelenk ▪ Wurfarm unterhalb des Schultergelenks eingedreht ▪ Speerhaltung zu locker, Finger nicht auf der Wicklung, Abkippen des Handgelenks ▪ Wurfbewegung beinhaltet ein Abbeugen im Hüftgelenk, wobei die Hüfte nicht nach vorne gebracht werden kann ▪ keine Wurfverzögerung, zu früher Armeinsatz ▪ kein Einsatz der Unterarmschleuder, Wurf mit weitestgehend gestrecktem Arm 	<p>Killing 2011, Weiler, Bauersfeld, Kunz</p>
<p>38. Speer mit zu großem Verkantungswinkel abgeworfen (größer als 10 Grad) <u>>> Mögliche Ursache:</u></p> <ul style="list-style-type: none"> ▪ Speer wird nach Rückführung (nach hinten seitlich) zum Abwurf hin nicht ausreichend in Wurfrichtung gebracht ▪ Wurfarm nicht richtig eingedreht 	<p>Killing 2011</p>

Tab. 49: Erwartungshorizont für das Fehlerwissen der Phase Abfangen

4. Fehlerbilder Abfangen	Autor
A) Übertreten	
<p>1. Übertreten der Abwurfmarkierung <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ zu hohe Anlaufgeschwindigkeit ▪ Abwurfstelle zu dicht an der Abwurfmarkierung ▪ ungenügende Bremswirkung des Stemmbeins ▪ mangelhafte Konzentration während der Abfangbewegung ▪ Anlauf unbewusst um zwei Schritte verlängert (bei Anfänger häufig) 	<p>Killing 2011, Strüder, Kunz</p>
B) Verschenken	
<p>2. Verschenken beim Abwurf (Abfangbewegung endet deutlich vor dem Abwurfbogen) <u>>> Mögliche Ursachen:</u></p> <ul style="list-style-type: none"> ▪ unpassende Anlaufmarke ▪ verkürzte Schrittlänge durch zu geringe Anlaufgeschwindigkeit ▪ mangelndes Arbeiten über das Stemmbein hinweg durch mangelnde Anlaufgeschwindigkeit oder passive Druckbearbeitung 	<p>Strüder</p>

<p>Test Recognition 1 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Machen Sie einen Schritt mit dem rechten Fuß um ca. 50 cm nach vorne. Nehmen Sie ihren linken Arm gerade nach oben. Bringen Sie ihren rechten Arm zur rechten Seite in die Waagrechte. Drehen Sie ihren Oberkörper um 90° zur rechten Seite. Drehen Sie ihren Kopf um 90° zur rechten Seite.</p>	<p>Test Recognition 6 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm zur linken Seite. Machen Sie einen Schritt mit dem rechten Fuß um ca. 50 cm nach vorne. Nehmen Sie das Kinn auf die Brust. Bringen Sie ihren linken Arm zur linken Seite in die Waagrechte. Bringen Sie ihren rechten Arm zur rechten Seite in die Waagrechte.</p>
<p>Test Recognition 2 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm nach hinten. Nehmen Sie ihren rechten Arm gerade nach oben. Drehen Sie ihren Oberkörper um 90° zur linken Seite. Nehmen Sie das Kinn auf die Brust. Bringen Sie ihren linken Arm zur linken Seite in die Waagrechte.</p>	<p>Test Recognition 7 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Nehmen Sie ihren rechten Arm nach oben. Drehen Sie ihren Kopf um 90° zur rechten Seite. Bringen Sie ihren linken Arm nach vorne in die Waagrechte. Drehen Sie ihren Oberkörper um 90° zur rechten Seite. Bringen Sie ihren rechten Arm zur rechten Seite in die Waagrechte.</p>
<p>Test Recognition 3 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Machen Sie einen Schritt mit dem rechten Fuß um ca. 50 cm nach hinten. Bringen Sie ihren linken Arm nach vorne in die Waagrechte. Nehmen Sie ihren rechten Arm gerade nach oben. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm zur linken Seite. Drehen Sie ihren Oberkörper um 90° zur rechten Seite.</p>	<p>Test Recognition 8 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Machen Sie einen Schritt mit dem rechten Fuß um ca. 50 cm nach vorne. Nehmen Sie ihren linken Arm nach oben. Drehen Sie ihren Oberkörper um 90° zur linken Seite. Drehen Sie ihren Kopf um 90° zur rechten Seite. Bringen Sie ihren rechten Arm zur rechten Seite in die Waagrechte.</p>
<p>Test Recognition 4 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Bringen Sie ihren linken Arm zur linken Seite in die Waagrechte. Bringen Sie ihren rechten Arm zur rechten Seite in die Waagrechte. Nehmen Sie das Kinn auf die Brust. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm zur linken Seite. Drehen Sie ihren Oberkörper um 90° zur linken Seite.</p>	<p>Test Recognition 9 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm nach vorne. Nehmen Sie das Kinn auf die Brust. Nehmen Sie ihren rechten Arm nach oben. Bringen Sie ihren linken Arm nach vorne in die Waagrechte. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm nach hinten.</p>
<p>Test Recognition 5 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Machen Sie einen Schritt mit dem rechten Fuß um ca. 50 cm zur rechten Seite. Bringen Sie ihren rechten Arm zur rechten Seite in die Waagrechte. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm zur linken Seite. Drehen Sie ihren Oberkörper um 90° zur rechten Seite. Nehmen Sie ihren linken Arm nach oben.</p>	<p>Test Recognition 10 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Bringen Sie ihren linken Arm nach vorne in die Waagrechte. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm zur linken Seite. Drehen Sie ihren Kopf um 90° zur rechten Seite. Machen Sie einen Schritt mit dem rechten Fuß um ca. 50 cm nach vorne. Bringen Sie ihren rechten Arm zur rechten Seite in die Waagrechte.</p>

Abb. 114: Anweisungen der zehn Aufgaben des TKBV (Teil 1)

<p>Test Regeneration 1 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm nach vorne. Nehmen Sie das Kinn auf die Brust. Machen Sie einen Schritt mit dem rechten Fuß um ca. 50 cm zur rechten Seite. Drehen Sie ihren Oberkörper um 90° zur rechten Seite. Bringen Sie ihren linken Arm zur linken Seite in die Waagrechte.</p>	<p>Test Regeneration 6 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Nehmen Sie ihren rechten Arm gerade nach oben. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm nach vorne. Drehen Sie ihren Kopf um 90° zur rechten Seite. Bringen Sie ihren linken Arm nach vorne in die Waagrechte. Machen Sie einen Schritt mit dem rechten Fuß um ca. 50 cm zur rechten Seite.</p>
<p>Test Regeneration 2 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Machen Sie einen Schritt mit dem rechten Fuß um ca. 50 cm nach vorne. Drehen Sie ihren Oberkörper um 90° zur linken Seite. Drehen Sie ihren Kopf um 90° zur rechten Seite. Bringen Sie ihren rechten Arm nach vorne in die Waagrechte. Bringen Sie ihren linken Arm nach oben in die Senkrechte.</p>	<p>Test Regeneration 7 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Nehmen Sie das Kinn auf die Brust. Machen Sie einen Schritt mit dem rechten Fuß um ca. 50 cm nach vorne. Bringen Sie ihren linken Arm nach vorne in die Waagrechte. Bringen Sie ihren rechten Arm zur rechten Seite in die Waagrechte. Drehen Sie ihren Oberkörper um 90° zur linken Seite.</p>
<p>Test Regeneration 3 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm nach hinten. Drehen Sie ihren Kopf um 90° zur linken Seite. Bringen Sie ihren linken Arm zur linken Seite in die Waagrechte. Drehen Sie ihren Oberkörper um 90° zur linken Seite. Bringen Sie ihren rechten Arm zur rechten Seite in die Waagrechte.</p>	<p>Test Regeneration 8 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm zur linken Seite. Drehen Sie ihren Oberkörper um 90° zur linken Seite. Bringen Sie ihren rechten Arm zur rechten Seite in die Waagrechte. Drehen Sie ihren Kopf um 90° zur rechten Seite. Nehmen Sie ihren linken Arm gerade nach oben.</p>
<p>Test Regeneration 4 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Bringen Sie ihren linken Arm nach vorne in die Waagrechte. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm nach vorne. Bringen Sie ihren rechten Arm nach oben in die Senkrechte. Drehen Sie ihren Kopf um 90° zur rechten Seite. Drehen Sie ihren Oberkörper um 90° zur rechten Seite.</p>	<p>Test Regeneration 9 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Nehmen Sie ihren rechten Arm gerade nach oben. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm nach vorne. Nehmen Sie das Kinn auf die Brust. Bringen Sie ihren linken Arm nach vorne in die Waagrechte. Drehen Sie ihren Oberkörper um 90° zur rechten Seite.</p>
<p>Test Regeneration 5 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Machen Sie einen Schritt mit dem rechten Fuß um ca. 50 cm zur rechten Seite. Drehen Sie ihren Oberkörper um 90° zur linken Seite. Bringen Sie ihren rechten Arm nach vorne in die Waagrechte. Bringen Sie ihren linken Arm zur linken Seite in die Waagrechte. Nehmen Sie das Kinn auf die Brust.</p>	<p>Test Regeneration 10 Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm nach hinten. Bringen Sie ihren rechten Arm zur rechten Seite in die Waagrechte. Nehmen Sie das Kinn auf die Brust. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm zur linken Seite. Bringen Sie ihren linken Arm zur linken Seite in die Waagrechte.</p>

Abb. 115: Anweisungen der zehn Aufgaben des TKBV (Teil 2)



Liebe Probanden,
 Markieren Sie bitte bei den Multiple-Choice-Fragen Ihre Antwort(en) in der folgenden Weise: ○ ⊗ ○ .
 Zur Korrektur einer bereits markierte Antwort füllen Sie diese vollständig aus, ungefähr so: ○ ⊕ ⊗ .
 Sollten Sie ein drittes Mal bei einer Antwortmöglichkeit Ihre Meinung ändern, so kreisen Sie das Auswahlfeld ein.
 Vielen Dank für Ihre Teilnahme!

Matrikelnummer

Bitte schreiben Sie unter diesen Text Ihren Namen und kreuzen Sie im rechten Bereich Ihre Matrikelnummer an (erste Ziffer = oberste Zeile).	0	1	2	3	4	5	6	7	8	9
	<input type="checkbox"/>									
	<input type="checkbox"/>									
	<input type="checkbox"/>									
	<input type="checkbox"/>									

TKBV

Wiedererkennung (Recognition)

1.1 Aufgabe 1: Entscheiden Sie, welches Bild Ihrer Bewegungsvorstellung entspricht.	<input type="checkbox"/> Bild 1 <input type="checkbox"/> Bild 2 <input type="checkbox"/> Bild 3 <input type="checkbox"/> Bild 4 <input type="checkbox"/> Bild 5
1.2 Beantwortungszeit für Aufgabe 1:	
1.3 Aufgabe 2: Entscheiden Sie, welches Bild Ihrer Bewegungsvorstellung entspricht.	<input type="checkbox"/> Bild 1 <input type="checkbox"/> Bild 2 <input type="checkbox"/> Bild 3 <input type="checkbox"/> Bild 4 <input type="checkbox"/> Bild 5
1.4 Beantwortungszeit für Aufgabe 2:	
1.5 Aufgabe 3: Entscheiden Sie, welches Bild Ihrer Bewegungsvorstellung entspricht.	<input type="checkbox"/> Bild 1 <input type="checkbox"/> Bild 2 <input type="checkbox"/> Bild 3 <input type="checkbox"/> Bild 4 <input type="checkbox"/> Bild 5
1.6 Beantwortungszeit für Aufgabe 3:	
1.7 Aufgabe 4: Entscheiden Sie, welches Bild Ihrer Bewegungsvorstellung entspricht.	<input type="checkbox"/> Bild 1 <input type="checkbox"/> Bild 2 <input type="checkbox"/> Bild 3 <input type="checkbox"/> Bild 4 <input type="checkbox"/> Bild 5
1.8 Beantwortungszeit für Aufgabe 4:	
1.9 Aufgabe 5: Entscheiden Sie, welches Bild Ihrer Bewegungsvorstellung entspricht.	<input type="checkbox"/> Bild 1 <input type="checkbox"/> Bild 2 <input type="checkbox"/> Bild 3 <input type="checkbox"/> Bild 4 <input type="checkbox"/> Bild 5
1.10 Beantwortungszeit für Aufgabe 5:	
1.11 Aufgabe 6: Entscheiden Sie, welches Bild Ihrer Bewegungsvorstellung entspricht.	<input type="checkbox"/> Bild 1 <input type="checkbox"/> Bild 2 <input type="checkbox"/> Bild 3 <input type="checkbox"/> Bild 4 <input type="checkbox"/> Bild 5
1.12 Beantwortungszeit für Aufgabe 6:	
1.13 Aufgabe 7: Entscheiden Sie, welches Bild Ihrer Bewegungsvorstellung entspricht.	<input type="checkbox"/> Bild 1 <input type="checkbox"/> Bild 2 <input type="checkbox"/> Bild 3 <input type="checkbox"/> Bild 4 <input type="checkbox"/> Bild 5
1.14 Beantwortungszeit für Aufgabe 7:	
1.15 Aufgabe 8: Entscheiden Sie, welches Bild Ihrer Bewegungsvorstellung entspricht.	<input type="checkbox"/> Bild 1 <input type="checkbox"/> Bild 2 <input type="checkbox"/> Bild 3 <input type="checkbox"/> Bild 4 <input type="checkbox"/> Bild 5
1.16 Beantwortungszeit für Aufgabe 8:	
1.17 Aufgabe 9: Entscheiden Sie, welches Bild Ihrer Bewegungsvorstellung entspricht.	<input type="checkbox"/> Bild 1 <input type="checkbox"/> Bild 2 <input type="checkbox"/> Bild 3 <input type="checkbox"/> Bild 4 <input type="checkbox"/> Bild 5
1.18 Beantwortungszeit für Aufgabe 9:	



1.19 Aufgabe 10: Entscheiden Sie, welches Bild Ihrer Bewegungsvorstellung entspricht.	<input type="radio"/> Bild 1 <input type="radio"/> Bild 2 <input type="radio"/> Bild 3 <input type="radio"/> Bild 4 <input type="radio"/> Bild 5
1.20 Beantwortungszeit für Aufgabe 10:	
1.21 Welche Strategie haben Sie bei der Vorstellungsbildung angewandt?	<input type="checkbox"/> 1-Person-Perspektive <input type="checkbox"/> 3-Person-Perspektive Fremd <input type="checkbox"/> 3-Person-Perspektive Selbst

Bewegungsvergleich

2.1 Aufgabe 1: Kreuzen Sie an, worin sich Bewegung 2 von Bewegung 1 unterscheidet.	<input type="checkbox"/> Reihenfolge der Einzelbewegungen stimmt nicht <input type="checkbox"/> Fehlerhafte Beinbewegung <input type="checkbox"/> Fehlerhafte Armbewegung <input type="checkbox"/> Fehlerhafte Oberkörperbewegung <input type="checkbox"/> Fehlerhafte Kopfbewegung
2.2 Bearbeitungszeit für Aufgabe 1:	
2.3 Aufgabe 2: Kreuzen Sie an, worin sich Bewegung 2 von Bewegung 1 unterscheidet.	<input type="checkbox"/> Reihenfolge der Einzelbewegungen stimmt nicht <input type="checkbox"/> Fehlerhafte Beinbewegung <input type="checkbox"/> Fehlerhafte Armbewegung <input type="checkbox"/> Fehlerhafte Oberkörperbewegung <input type="checkbox"/> Fehlerhafte Kopfbewegung
2.4 Bearbeitungszeit für Aufgabe 2:	
2.5 Aufgabe 3: Kreuzen Sie an, worin sich Bewegung 2 von Bewegung 1 unterscheidet.	<input type="checkbox"/> Reihenfolge der Einzelbewegungen stimmt nicht <input type="checkbox"/> Fehlerhafte Beinbewegung <input type="checkbox"/> Fehlerhafte Armbewegung <input type="checkbox"/> Fehlerhafte Oberkörperbewegung <input type="checkbox"/> Fehlerhafte Sprungbewegung
2.6 Bearbeitungszeit für Aufgabe 3:	
2.7 Aufgabe 4: Kreuzen Sie an, worin sich Bewegung 2 von Bewegung 1 unterscheidet.	<input type="checkbox"/> Reihenfolge der Einzelbewegungen stimmt nicht <input type="checkbox"/> Fehlerhafte Beinbewegung <input type="checkbox"/> Fehlerhafte Armbewegung <input type="checkbox"/> Fehlerhafte Oberkörperbewegung <input type="checkbox"/> Fehlerhafte Kopfbewegung
2.8 Bearbeitungszeit für Aufgabe 4:	
2.9 Aufgabe 5: Kreuzen Sie an, worin sich Bewegung 2 von Bewegung 1 unterscheidet.	<input type="checkbox"/> Reihenfolge der Einzelbewegungen stimmt nicht <input type="checkbox"/> Fehlerhafte Armbewegung <input type="checkbox"/> Fehlerhafte Oberkörperbewegung <input type="checkbox"/> Fehlerhafte Kopfbewegung <input type="checkbox"/> Gleiche Endposition nach Bewegungsausführung
2.10 Bearbeitungszeit für Aufgabe 5:	



<p>2.11 Aufgabe 6: Kreuzen Sie an, worin sich Bewegung 2 von Bewegung 1 unterscheidet.</p>	<input type="checkbox"/> Reihenfolge der Einzelbewegungen stimmt nicht <input type="checkbox"/> Fehlerhafte Beinbewegung <input type="checkbox"/> Fehlerhafte Armbewegung <input type="checkbox"/> Fehlerhafte Kopfbewegung <input type="checkbox"/> Fehlerhafte Sprungbewegung
<p>2.12 Bearbeitungszeit für Aufgabe 6:</p>	
<p>2.13 Aufgabe 7: Kreuzen Sie an, worin sich Bewegung 2 von Bewegung 1 unterscheidet.</p>	<input type="checkbox"/> Reihenfolge der Einzelbewegungen stimmt nicht <input type="checkbox"/> Fehlerhafte Beinbewegung <input type="checkbox"/> Fehlerhafte Armbewegung <input type="checkbox"/> Fehlerhafte Kopfbewegung <input type="checkbox"/> Gleiche Endposition nach Bewegungsausführung
<p>2.14 Bearbeitungszeit für Aufgabe 7:</p>	
<p>2.15 Aufgabe 8: Kreuzen Sie an, worin sich Bewegung 2 von Bewegung 1 unterscheidet.</p>	<input type="checkbox"/> Reihenfolge der Einzelbewegungen stimmt nicht <input type="checkbox"/> Fehlerhafte Beinbewegung <input type="checkbox"/> Fehlerhafte Armbewegung <input type="checkbox"/> Fehlerhafte Oberkörperbewegung <input type="checkbox"/> Fehlerhafte Sprungbewegung
<p>2.16 Bearbeitungszeit für Aufgabe 8:</p>	
<p>2.17 Aufgabe 9: Kreuzen Sie an, worin sich Bewegung 2 von Bewegung 1 unterscheidet.</p>	<input type="checkbox"/> Reihenfolge der Einzelbewegungen stimmt nicht <input type="checkbox"/> Fehlerhafte Armbewegung <input type="checkbox"/> Fehlerhafte Oberkörperbewegung <input type="checkbox"/> Fehlerhafte Kopfbewegung <input type="checkbox"/> Gleiche Endposition nach Bewegungsausführung
<p>2.18 Bearbeitungszeit für Aufgabe 9:</p>	
<p>2.19 Aufgabe 10: Kreuzen Sie an, worin sich Bewegung 2 von Bewegung 1 unterscheidet.</p>	<input type="checkbox"/> Reihenfolge der Einzelbewegungen stimmt nicht <input type="checkbox"/> Fehlerhafte Armbewegung <input type="checkbox"/> Fehlerhafte Oberkörperbewegung <input type="checkbox"/> Fehlerhafte Kopfbewegung <input type="checkbox"/> Fehlerhafte Sprungbewegung
<p>2.20 Bearbeitungszeit für Aufgabe 10:</p>	

Abb. 116: Testbogen für den TKBV-Teil 1 und den Bewegungsbeobachtungstest

Kontrollierbarkeit von Bewegungsvorstellungen – Wiedererkennen (recognition)

Untersuchungsziel

Ziel der Untersuchung ist es, die Kontrollierbarkeit mental vorgestellter Bewegungen zu überprüfen.

Untersuchungsablauf

Die Untersuchungen finden jeweils einzeln in der Turnhalle (oder einem anderen geeigneten Raum) statt, damit keine Einwirkungen von Umweltfaktoren wie z.B. Zuschauer oder Hintergrundgeräusche stattfinden. Nach der Begrüßung werden die persönlichen Daten des Untersuchungsteilnehmers (Name, Alter, Geschlecht, sportliche Aktivität) festgehalten. Im Anschluss daran erhält der Teilnehmer eine allgemeine Instruktion zu der geplanten Untersuchung. Auftretende Fragen werden direkt beantwortet, so dass sich die Versuchsperson völlig auf die weiteren Anweisungen konzentrieren kann.

Allgemeine Instruktionen

„(Vielen Dank, dass Sie bei unserem Experiment teilnehmen). Sie werden ein Set von 6 aufeinander folgenden Instruktionen bekommen, welche sie mental nachvollziehen müssen. Pro Instruktion wird immer nur ein Körperteil bewegt (Kopf, Arme, Beine, Rumpf). Aus einem Set von 5 Bildern müssen sie so schnell wie möglich die Endposition auswählen, die ihrer Vorstellung am nächsten kommt. Wichtig ist, dass Sie sehr konzentriert mitmachen. Davon hängt die Güte unserer Ergebnisse ab. Versuchen Sie sich immer möglichst eng an die Instruktionen zu halten.“

Spezifische Instruktionen

„Schließen Sie die Augen, stellen sie sich vor ihren Körper zu bewegen ohne tatsächlich eine aktive Bewegung auszuführen, während sie den Instruktionen folgen. Ich gebe Ihnen nun 6 aufeinanderfolgende Instruktionen.“

Beispielstest:

1. Nehmen Sie eine gerade Haltung mit geschlossenen Füßen ein.
2. Machen Sie einen Schritt mit dem linken Fuß um ca. 50 cm nach hinten.
3. Drehen Sie ihren Kopf um 90° zur linken Seite.
4. Bringen Sie ihren linken Arm zur linken Seite in die Waagrechte.
5. Drehen Sie ihren Oberkörper um 90° zur linken Seite.
6. Nehmen Sie ihren rechten Arm gerade nach oben.

„Öffnen Sie nun die Augen und wählen sie so schnell wie möglich aus diesen 5 Bildern dasjenige aus, das ihrer Vorstellung am nächsten kommt.“

Nach dieser Vorstellung der spezifischen Instruktionen werden die Untersuchungsteilnehmer gefragt, ob es noch Unklarheiten gibt oder Fragen zum Versuchsablauf; ist dies nicht der Fall, wird mit der Untersuchung angefangen.

Abb. 117: Auszug aus dem original TKBV-Testmanual von Schott (Seite 1)

Tab. 50: Beobachtungsdetails der Wurfbewegung (B.Mw = Nummerierung des biomechanischen Messwerts aus Tab. 51)

		QS 1	QS2	QS3	B.Mw
Impulsschritt	1	Wurfarm gestreckt in Schulterhöhe		Stark gebeugt	
	2	Rumpf zurückgeneigt und verwrungen		Vorgeneigt	#8
	3	Schwungbein: Knieführung, Fuß vorge-dreht		Fuß weit vorne	
	4	Abdruckbein: flacher Impuls		Hohe Flugphase	#9
Wurfauslage	5	Wurfarm: unverändert auf Schulterhöhe		Deutlich unter Schulter-höhe	#10
	6	Wurfarm: gestreckt		Stark gebeugt	#11 #12
	7	Rumpf: deutliche Rücklage		Vorlage	#13
	8	Druckfußaufsatz: Fußspitze in Wurfrich-tung		Fußspitze zeigt quer oder zurück	
	9	Körperschwerpunkt: abgesenkt hinter Druckbein		Druckbein relativ ge-streckt	
	10	Stemmbein: bereits vor Druckbein		Deutlich hinter Druckbein	#14
Stemmschritt	11	Wurfarm: lang		Stark gebeugt	#16 #17
	12	Wurfarm: leicht angehoben		Deutlich unter Schulter-höhe	#15
	13	Schwungarm: lang, Seitrückriss		Bereits neben Körper	
	14	Schulterachse: in Wurfrichtung		Bereits frontal	
	15	Druckbeinknie: horizontale Schubbe-schleunigung		Passiv ohne Druck	
	16	Stemmbein: gestreckt		Stark gebeugt	#19
	17	Stemmbein: über die Ferse gesetzt		Auf dem Ballen	#20
Bogenspannung	18	Wurfarm: angebeugt auf Schulterhöhe		Relativ gestreckt und tief	
	19	Wurfarmseite: explosiver Schwenk in Wurfrichtung		Ohne erkennbare Be-schleunigung	
	20	Brust/Becken: frontal in Wurfrichtung		Parallel zur Wurfrichtung	
	21	Stemmbein: Ganzsohlenkontakt, ge-streckt		Fußspitze oder ohne Bo-denkontakt	
	22	Druckbein: noch Bodenkontakt		Kein Bodenkontakt	
	23	Schwungarm: gebeugt und fixiert		Weit hinter Körper	
Schlagposition	24	Wurfarm: Ellbogenführung über Schulter		Ellbogen	
	25	Stemmbein: gestreckt, Ganzsohlenkon-takt		Stark gebeugt	#21
	26	Druckbein: Schleifkontakt		Ohne Bodenkontakt	
	27	Wurfarm: Nachschlagen in Wurfrichtung		Ohne Nachschlagen	#22 #23
	28	Becken: Abbremsen (Knick)		Gestreckter Körper	
Abfangen	29	Rumpf: nach vorne in Wurfrichtung ge-neigt		Oberkörperrücklage	#24
	30	Stemmbein: gestreckt, Hebel im Ballen-stand		KSP bereits vor dem Stemmbein	

Testmanual

Manual Biomechanische Analyse

11 = Wert muss in die Excel-Tabelle in Spalte 1 eingetragen werden

Grundsätzliches:

- für alle Marker, Linien, Winkel etc. die Strichbreite 1 verwenden (erhöht die spätere Messgenauigkeit!)

Analyseschritt-1

1. Zeitliche Gestaltung der Bodenkontakte bestimmen

- Messzeitpunkte:
 - t4 = viertletzter Bodenkontakt
 - t3 = drittletzter Bodenkontakt
 - t2 = Druckbein setzt auf
 - t1 = Stemmbein setzt auf
 - t0 = Wurfgerät verlässt die Wurfhand
- Schritte:
 - L3 = drittletzter Schritt
 - L2 = Impulsschritt
 - L1 = Stemmschritt
- Dartfish: Flugzeiten bestimmen für L3 (= t4-t3), L2 (= t3-t2) **17** und L1 (= t2-t1) **18**
 - immer bei der ersten Berührung des entsprechenden Bodenkontakts (egal ob Ferse oder Fußspitze) >> **Zeitmessung** einfügen (rechte Maustaste: Timecode auf Null setzen > Zählen beginnen)
 - Bodenkontakt (immer Fußspitze!) mit **Marker** markieren
 - Bis zum nächsten Bodenkontakt vortrimmen (Strg + Pfeiltaste) >> rechte Maustaste auf Zeitmessung >> Zählen beenden
 - neue Zeitmessung einfügen (Timecode auf Null setzen > Zählen beginnen)
 - Bodenkontakt (Fußspitze) mit Marker markieren ... usw.
 - Beachte: bei Zeitlupenaufnahmen Zeitwert durch 4 teilen (Slowmotion = 4-fach verlangsamt), um den Wert in Echtzeit zu errechnen

2. Räumliche Gestaltung der Bodenkontakte bestimmen

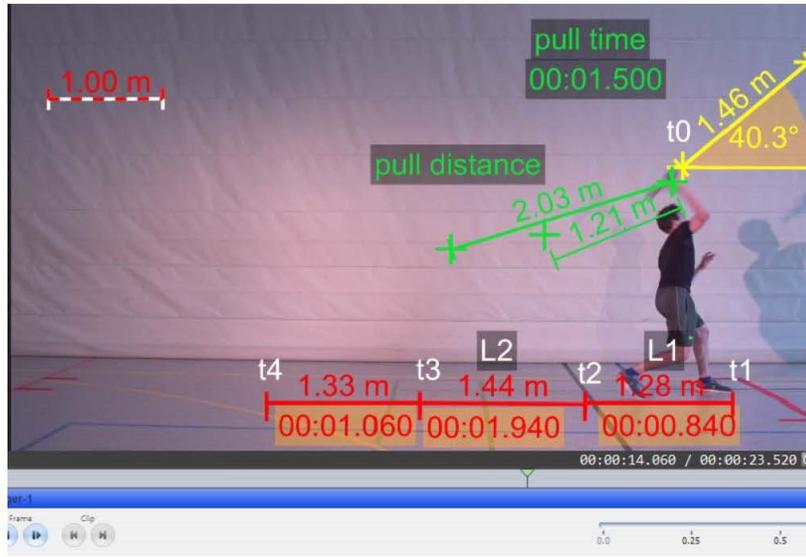
- Dartfish: Schrittlänge bestimmen für L3 **11**, L2 **2**, L1 **3**
 - **Entfernungsmessung** kalibrieren
 - Normmeter einfügen (erscheint automatisch bei erster Benutzung des Entfernungstools) >> neue Entfernungsmessung auf grüne Linien des Badmintonfeldes zwischen Grundlinie und Aufschlaglinie-Doppel als Referenzgröße einstellen (= 76 cm >> Normmeter verstellen, bis 76 cm als Wert zwischen den grünen Linien erscheinen)
 - **Entfernungstool** einfügen >> auf Distanz zwischen zwei Markern einstellen >> Schrittlänge ermitteln

3. Abwurfzeit (= pull time) und Beschleunigungsweg (= pull distance) messen

- dafür neue **Zeitmessung** (oben rechts) einfügen
 - Beginn: erste Bodenberührung des Druckbeins (= t2) >> an diesem Moment auch gleich den Marker auf die Wurfhand (Messpunkt: Übergang Handfläche - Finger) setzen
 - Ende: Wurfgerät verlässt die Wurfhand (= t0) **13** >> auch gleich wieder den Marker auf die Wurfhand setzen
- **Entfernungsmessung** zwischen den beiden Markern = pull distance (t2-t0) **14**
- Pull distance (t1-t0) **15**
 - Marker auf Fingerspitzen bei t1 (= Stemmbein berührt Boden) setzen
 - Entfernung zwischen den beiden Markern messen

4. Abflugwinkel und Abfluggeschwindigkeit ermitteln

- Bei t0: Spitze des Wurfgeräts mit Marker markieren
 - 10 Frames weitertrimmen (oder auch weniger Frames, wenn das Wurfgerät sonst aus dem Bildschirm verschwindet) >> Wurfgerät (Spitze) mit Marker markieren
 - **Abwurfwinkel** bestimmen (Schenkel 1 des Winkeltools parallel zur Linie des Trennvorhangs, Schenkel 2 von Marker zu Marker) [26]
 - Flugweite des Wurfgeräts messen (**Entfernungstool** zwischen den beiden Markern einfügen) [27]
 - Aus Flugweite und Zeit (= 10 Frames bei t=0,005s pro Frame) >> Geschwindigkeit berechnen $v = s/t = \text{Flugweite} / 0,05s$ >> Beispiel unten: $v = 1,46/0,05 = 29,2 \text{ m/s}$



Analyseschritt-2

5. Anlaufgeschwindigkeit

- Messbar über Strecke des Körperschwerpunkts
 - Messzeitpunkte:
 - Bodenberührung drittletzter Kontakt >> KSP (T-Shirt-Kragen, Aufdruck, evtl. auch Ohrläppchen) markieren >> 10 Frames zurücktrimmen >> **Entfernung** messen >> Geschwindigkeit in m/s = Wert in m x 20 [10]
 - Bodenberührung vorletzter Kontakt >> wieder 10 Frames rückwärts >> Entfernung messen >> v berechnen [11]
 - Bodenberührung letzter Kontakt >> wieder 10 Frames rückwärts >> Entfernung messen >> v berechnen [12]



Analyseschritt-3

6. Ellbogenwinkel

- Winkeltool:** Gelenkpunkt Ellbogen als Anfangspunkt >> Schenkel auf die Außenseite Oberarm und Unterarm ziehen
 - Im Moment des Aufsetzens des Druckbeins (t2) 24
 - Im Moment des Aufsetzens des Stemmbeins (t1) 25
 - Im Moment des Abwurfs (t0) 26

7. Stemmbeinwinkel

- Winkeltool:** Gelenkpunkt Knie Scheibe >> Schenkel auf Vorderseite Oberschenkel und Vorderseite Unterschenkel
 - Im Moment des Aufsetzens des Stemmbeins (t1) 22

Analyseschritt-4

8. Oberkörperpositionen

- Winkelmessung:** Vertikale (mit Viereck als Vorgabe für die Senkrechte) >> Winkeltool so anlegen, dass der Gelenkpunkt am Hüftgelenk anliegt >> ein Schenkel senkrecht nach oben, der andere auf der Linie Hüfte-Schulter
 - Rücklage = negativer Wert (siehe Bild bei t2 = -17,2°)
- Messpunkte:**
 - t2 >> 5
 - t1 >> 6
 - t0 >> 21

9. Abwurfhöhe

- Entfernung messen: zwischen Bodenmarker t1 und Handpositionsmarker t0 27

10. Stemmbeinwinkel

- Gelenkpunkt:** Knie Scheibe >> Schenkel auf Vorderseite Oberschenkel und Vorderseite Unterschenkel
 - Messpunkt: t0 23

Analyseschritt-5

11. Acceleration path

- Entfernung zwischen Wurfhand und hinterer Hüfte (Wurfarmseite) >> über eingezeichnetes Viereck messbar (siehe Bild unten)
- Messpunkte:
 - t2 18
 - t1 19

12. Wurfarmhöhe

- Entfernung zwischen Wurfhand und Schulterhöhe (Wurfarmseite) >> von Viereck-Oberkante (t2) bis Schulteroberkante
- Messpunkt:
 - t2 20

Abb. 118: Testmanual Biomechanisch Analyse Wurfbewegung mit Dartfish

Tab. 51: Biomechanische Messwerte der Dartfish-Analyse

	#	Beschreibung	QS 1	QS2	QS3
Anlauf	1	Anlaufgeschwindigkeit vor t_3 (in m/s)			
	2	Anlaufgeschwindigkeit vor t_2 (in m/s)			
	3	Anlaufgeschwindigkeit vor t_1 (in m/s)	$\geq 3,4$ m/s		$\leq 2,6$ m/s
	4	Schrittlänge von L2			
	5	Zeitintervall zwischen t_3 und t_2			
	6	Schrittlänge von L1			
	7	Zeitintervall zwischen t_2 und t_1			
			Zeitlicher Quotient L1/L2	$\leq 0,56$	
		Räumlicher Quotient L1/L2	$\leq 0,88$		$\geq 1,22$
Impulsschritt	8	Rumpfwinkel zum Zeitpunkt $t_{3\text{Ende}}$	$\leq -9,5^\circ$		$\geq -2,4^\circ$
	9	Höhe der Körperschwerpunktkurve während des Impulsschrittes			
Wurfauslage	10	Höhe der Wurfhand im Vergleich zur Höhe des Schultergelenks in t_2	$-0,1\text{m} \leq x \leq 0,1\text{m}$	$-0,30\text{m} < x < -0,1\text{m}$ oder $x > 0,1\text{m}$	$\leq -0,30\text{m}$
	11	Abstand der Wurfhand von der wurfseitigen Hüfte in t_2	$\geq 0,60\text{m}$		$\leq 0,40\text{m}$
	12	Ellbogenwinkel in t_2	$\geq 150^\circ$		$\leq 105^\circ$
	13	Rumpfneigung in t_2	$\geq -15^\circ$		$\leq -6,5^\circ$
	14	Abstand Druckbeinfußspitze – Stemmbeinfußspitze in t_2	$\geq 0,1\text{m}$		$\leq -0,2\text{m}$
Stemmschritt	15	Höhe der Wurfhand im Vergleich zur Höhe des Schultergelenks in t_1	$0,05\text{m} \leq x \leq 0,2\text{m}$	$-0,1 < x < 0,05$ oder $x > 0,2\text{m}$	$x \leq -0,1\text{m}$
	16	Abstand der Wurfhand von der wurfseitigen Hüfte in t_1	$\geq 0,60\text{m}$		$\leq 0,40\text{m}$
	17	Ellbogenwinkel in t_1	$\geq 150^\circ$		$\leq 105^\circ$
	18	Höhe der Wurfhand im Vergleich zur Höhe des Schultergelenks in t_1			
	19	Stemmbeinkniewinkel in t_1	$\geq 165^\circ$		$\leq 150^\circ$
	20	Fußgelenkwinkel (zwischen Horizontale und Fußsohle)	$\geq 20^\circ$		$\leq -20^\circ$
Schlagposition	21	Stemmbeinkniewinkel in t_0	$\geq 170^\circ$		$\leq 160^\circ$
		Unterschied Stemmbeinkniewinkel $t_1 - t_0$	$\geq -10^\circ$		$\leq -30^\circ$
	22	Ellbogenwinkel in t_0			
	23	Winkel zwischen Ellbogen und Vertikallinie in t_0	$\leq 20^\circ$		$\geq 40^\circ$
	24	Rumpfneigung in t_0	$\geq 5^\circ$		$\leq -5^\circ$

	25	Abwurfhöhe in t_0			
	26	Abflugwinkel des Wurfgeräts	$35 \leq x \leq 45$	$x \geq 45^\circ$ oder $25 < x < 35$	$x \leq 25^\circ$
	27	Abfluggeschwindigkeit in t_0	≥ 25 m/s		≤ 20 m/s
Dynamische Parameter	28	Pull distance ($t_2 > t_0$)	$\geq 2,00$ m		$\leq 1,75$ m
	29	Pull distance ($t_1 > t_0$)			
	30	Pull time ($t_2 > t_0$)			

Tab. 52: Biomechanische Messdaten einer Testperson aus der Videoanalyse der Wurfbewegung

	Anlaufgestaltung								Anlaufgeschwindigkeit			
	1	2	3	4	5	6	7	8	9	10	11	12
	räumlich (in m)			Rumpfwinkel		zeitlich (in s)			horizontaler Anteil			
L2	L1	L1/L2	t2	t1	L2	L1	L1/L2	vor t3	vor t2	vor t1	t3>t1	
#31	1,35	1,10	1,23	-23,9	-27,6	0,000	0,000		0,00	0,00	0,00	

Wurfbeschleunigung						
13	14	15	16	17	18	19
Pull time in s ($t_2 > t_0$)	Pull distance ($t_2 > t_0$)		P.dist./P.time		Accel.path	
	($t_2 > t_0$)	t1-t0		t2	t1	Höhe W.h.t2
0	1,82	1,11		0,34	0,49	0,15

Abwurfposition					Abflugparameter	
20	21	22	23	24	25	26
Rumpfwinkel in ° in t_0	Stemmbeinwinkel in ° in t_1	in t_0	Ellbog. in ° in t_1	Abw.höhe	Abflugwinkel in °	Abfluggeschw. in m/s
14,5	168,9	151,7	95,3	1,69	24,4	22,0

Tab. 53: Beurteilungsraster für die Qualitätsstufen der Bewegungsmerkmale

Bewegungsmerkmale	Qualitätsstufen		
	QS 1: Ideale Ausführung Merkmal entspricht Sollwert	QS 2: Geringe Abweichung Merkmal in Grundzügen erfüllt	QS 3: Starke Abweichung Merkmal kaum erfüllt
Bewegungsharmonie <i>der Gesamtausführung</i>	<ul style="list-style-type: none"> flüssig harmonisch koordiniert 	<ul style="list-style-type: none"> Gesamtkoordination vorhanden vereinzelte Mängel erkennbar 	<ul style="list-style-type: none"> eckig/unrund unkoordiniert falsches Stemmbein vorne
Bewegungstempo <i>des Anlaufs</i>	<ul style="list-style-type: none"> hohes Endtempo dynamisch (steigernd) 	<ul style="list-style-type: none"> mittleres Lauftempo undynamisch (konstant) 	<ul style="list-style-type: none"> Angehen Tempo stark abnehmend
Bewegungsrhythmus <i>der letzten beiden Schritte</i>	<ul style="list-style-type: none"> Lang-kurz-kurz mit richtiger Schrittfolge 	<ul style="list-style-type: none"> Hopser auf den letzten beiden Schritten 	<ul style="list-style-type: none"> keine Rhythmusänderung Tippelschritte vor dem Abwurf
Bewegungsumfang <i>der Ballrückführung in Wurfauslage</i>	<ul style="list-style-type: none"> Schulter-/Hüftachse parallel zur Wurfrichtung zurückgedreht Wurfarm nahezu gestreckt 	<ul style="list-style-type: none"> Beschleunigungsweg des Balles erkennbar reduziert (z. B. durch stark gebeugten Wurfarm) 	<ul style="list-style-type: none"> Oberkörper zeigt frontal in Wurfrichtung Falsches Stemmbein vorne
Bewegungskopplung <i>der Wurfbewegung</i>	<ul style="list-style-type: none"> Nacheinander von Hüfte/Schulter – Ellbogen – Wurfhand Bogenspannung erkennbar 	<ul style="list-style-type: none"> zu früher Ellbogeneinsatz Wurfarmschleuder erkennbar Bogenspannung schwach ausgeprägt 	<ul style="list-style-type: none"> Bogenspannung überhaupt nicht ausgeprägt (tiefer Ellbogen!) Abwurf über falsches Stemmbein zu früher Abwurf (noch vor dem Setzen des Stemmbeins)
Bewegungspräzision <i>des Abstoppens</i>	<ul style="list-style-type: none"> deutliches Abstoppen beim Abwurf gestrecktes Stemmbein 	<ul style="list-style-type: none"> geringes Abbremsen der Wurfbewegung stark gebeugtes Stemmbein 	<ul style="list-style-type: none"> Überlaufen der Wurfbewegung ohne Abstoppen Verlust des Gleichgewichtes



Lieber Proband,
 dieser Test wird maschinell ausgewertet. Markieren Sie bitte bei den Multiple-Choice-Fragen Ihre Antwort in der folgenden Weise: .
 Wenn Sie eine Antwort korrigieren möchten, füllen Sie bitte die falsch markierte Antwort vollständig aus, ungefähr so: .

Viel Dank für Ihre Teilnahme!

Matrikelnummer

Bitte schreiben Sie unter diesen Text Ihren Namen und kreuzen Sie im rechten Bereich Ihre Matrikelnummer an (erste Ziffer = oberste Zeile).		0	1	2	3	4	5	6	7	8	9
	_____	<input type="radio"/>									
	_____	<input type="radio"/>									
	_____	<input type="radio"/>									
	_____	<input type="radio"/>									
	_____	<input type="radio"/>									

Beurteilungstest Ballwurf

Zwischenzeitliche Beschäftigung mit der Schlagwurfbewegung

Haben Sie sich in der Zeit zwischen dem Fachwissenstest und dem heutigen Beurteilungstest mit der Theorie des Ballwurfs beschäftigt?
<input type="radio"/> ja <input type="radio"/> nein
Haben Sie in der Zeit zwischen dem Fachwissenstest und dem heutigen Beurteilungstest die Schlagwurfbewegung gezielt trainiert (z.B. durch Übungen, Kurs)?
<input type="radio"/> ja <input type="radio"/> nein

Vorerfahrung Bewegungsbeurteilung

Haben Sie Vorerfahrungen in der Beurteilung von leichtathletischen Wurfbewegungen (z.B. durch Trainerausbildung oder Trainertätigkeit)?
<input type="radio"/> ja <input type="radio"/> nein
Haben Sie Vorerfahrungen in der Bewegungsbeurteilung von dem leichtathletischen Schlagwurf ähnlichen Bewegungen (z.B. Handballschlagwurf, Volleyballaufschlag, Badmintonsmash)?
<input type="radio"/> ja <input type="radio"/> nein
Haben Sie Vorerfahrungen in der Beurteilung anderer sportlicher Bewegungen (z.B. durch Trainertätigkeiten in anderen Sportarten wie Fußball o.ä.), die nichts mit dem leichtathletischen Schlagwurf zu tun haben?
<input type="radio"/> ja <input type="radio"/> nein
Wenn ja, in welcher Sportarten bzw. welchen Sportarten?
Erreichte Punktzahl: <input type="radio"/> 0 <input type="radio"/> 1

Abb. 119: Testbogen zur Erfassung der Beurteilungsvorerfahrung

Tab. 54: Korrelationsmatrix zwischen der Beurteilungsleistung von Teil D und den möglichen Einflussfaktoren

Spearman-Rho		BT_D_ORM	FW_A_18_Items	FW_B_14_Items	Fehlerwissen	Bildreihe	Eigenrealisation	TKBV1	TKBV2	Vorerfahrung
BT_D_ORM	r	1,000	0,038	0,008	0,183	-0,174	,376**	-0,192	-0,039	0,103
	p		0,774	0,953	0,166	0,187	0,003	0,146	0,767	0,437
	N	59	59	59	59	59	59	59	59	59
FW_A_18_Items	r		1,000	,328*	0,220	0,154	0,234	0,209	0,119	-0,086
	p			0,011	0,094	0,243	0,075	0,112	0,370	0,519
	N		59	59	59	59	59	59	59	59
FW_B_14_Items	r			1,000	0,006	0,158	0,076	0,250	0,189	-0,004
	p				0,966	0,233	0,566	0,056	0,152	0,975
	N			59	59	59	59	59	59	59
Fehlerwissen	r				1,000	0,007	0,069	-0,057	0,017	-0,107
	p					0,959	0,605	0,669	0,898	0,419
	N				59	59	59	59	59	59
Bildreihe	r					1,000	-0,149	0,161	0,000	-0,239
	p						0,260	0,224	0,998	0,069
	N					59	59	59	59	59
Eigenrealisation	r						1,000	-0,069	0,227	0,133
	p							0,601	0,084	0,317
	N						59	59	59	59
TKBV1	r							1,000	0,226	-0,077
	p								0,085	0,560
	N							59	59	59
TKBV2	r								1,000	-0,003
	p									0,980
	N								59	59

Tab. 55: Korrelationsbeziehungen zwischen der Beurteilungsleistung von Teil B des Beurteilungstests und den möglichen Einflussfaktoren

Spearman-Rho		FW_A_18_Items	FW_B_14_Items	Fehlerwissen	Bildreihe	Eigenrealisation	TKBV 1	TKBV 2	Vorerfahrung
BT_B_ORM	r	0,007	0,081	0,025	0,031	-,487**	-0,019	0,045	-0,097
	p	0,956	0,541	0,849	0,813	0,000	0,885	0,736	0,463
	N	59	59	59	59	59	59	59	59
BT_B_DRM	r	0,129	0,027	0,171	0,173	-,430**	-0,037	0,078	-0,118
	p	0,332	0,837	0,194	0,189	0,001	0,780	0,557	0,375
	N	59	59	59	59	59	59	59	59

Tab. 56: Korrelationsbeziehungen zwischen der Beurteilungsleistung von Teil A des Beurteilungstests und den möglichen Einflussfaktoren

Spearman-Rho		FW_A_18_Items	FW_B_14_Items	Fehlerwissen	Bildreihe	Eigenrealisation	TKBV1	TKBV2	Vorerfahrung
BT_A_6	r	0,029	,290*	0,129	0,068	0,090	0,076	0,142	-0,023
	p	0,826	0,026	0,332	0,611	0,500	0,568	0,284	0,864
	N	59	59	59	59	59	59	59	59
BT_A_ORM	r	0,033	,278*	0,097	0,061	0,051	0,045	0,152	0,017
	p	0,807	0,033	0,465	0,644	0,700	0,738	0,249	0,899
	N	59	59	59	59	59	59	59	59

Tab. 57: Korrelationsbeziehungen zwischen der Beurteilungsleistung von Teil C des Beurteilungstests und den möglichen Einflussfaktoren

Spearman-Rho		FW_A_18_Items	FW_B_14_Items	Fehlerwissen	Bildreihe	Eigenrealisation	TKBV1	TKBV2	Vorerfahrung
BT_C_ORM	r	-0,157	-0,005	0,052	-0,206	0,185	-0,196	0,231	0,007
	p	0,235	0,972	0,694	0,118	0,160	0,137	0,079	0,957
	N	59	59	59	59	59	59	59	59

Tab. 58: Korrelationsbeziehungen zwischen der Beurteilungsleistung von Teil E des Beurteilungstests und den möglichen Einflussfaktoren

Spearman-Rho		FW_A_18_Items	FW_B_14_Items	Fehlerwissen	Bildreihe	Eigenrealisation	TKBV1	TKBV2	Vorerfahrung
BT_E	r	-0,076	-0,219	0,033	-0,207	0,186	-0,099	-0,018	0,045
	p	0,566	0,096	0,805	0,116	0,159	0,455	0,895	0,733
	N	59	59	59	59	59	59	59	59
BT_E_ORM	r	-0,129	-0,149	-0,017	-0,121	0,179	-0,144	0,035	0,071
	p	0,329	0,261	0,899	0,361	0,176	0,277	0,792	0,593
	N	59	59	59	59	59	59	59	59

Tab. 59: Likelihood-Ratio-Tabelle für Teil A und B des Fachwissenstests

	AIC	BIC	log.Lik	LRT	Df	p.value
Rasch (Teil A)	2061,16	2130,27	-997,58			
2PL-Modell (Teil A)	2058,58	2192,61	-965,29	64,58	31	< 0,001
Rasch (Teil B)	1188,62	1224,22	-577,31			
2PL-Modell (Teil B)	1186,75	1253,77	-561,38	31,87	15	0,007

Tab. 60: p-Werte der Teststatistik T4 für Teil A des Fachwissenstests

	A1.1.1	A1.1.2	A1.1.3	A1.1.4	A2.1.1	A2.1.2	A2.1.3	A2.1.4	A3.1.1	A3.1.2	A3.1.3	A3.1.4
p-Wert (gender)	0,975	0,376	0,774	0,890	0,997	0,684	0	0,567	0,832	0,696	0,689	0,329
p-Wert (mean Löser, N = 59)	0,842	0,881	0,510	0,785	0,349	0,780	0,950	0,399	0,883	0,248	0,981	0,049
Mean (Nicht-Löser, N = 59)	0,433	0,619	0,761	0,584	0,848	0,435	0,112	0,867	0,282	1,0	0,079	0,990

	A3.2.1	A3.2.2	A3.2.3	A3.2.4	A3.3.1	A3.3.2	A3.3.3	A3.3.4	A4.1.1	A4.1.2	A4.1.3	A4.1.4
p-Wert (gender)	0,938	0,049	0,580	0,283	0,727	0,569	0,704	0,402	0,784	0,172	0,894	0,500
p-Wert (mean)	0,678	0,435	0,503	0,427	0,135	0,100	0,596	0,906	0,754	0,939	0,924	0,678
p-Wert (mean Nicht-Löser, N = 59)	0,575	0,765	0,758	0,829	0,959	0,972	1,0	0,194	0,479	0,217	0,235	0,457

	A4.2.1	A4.2.2	A4.2.3	A4.2.4	A5.1.1	A5.1.2	A5.1.3	A5.1.4				
p-Wert (gender)	0,632	0,658	0,709	0,511	0,842	0,712	0,449	0,991				
p-Wert (mean)	0,681	0,609	0,707	0,277	0,991	0,774	0,180	0,542				
	0,491	1,0	0,511	0,941	0,051	0,536	0,919	0,643				

Tab. 61: p-Werte der Teststatistik T4 für Teil B des Fachwissenstests mit Splitkriterium Geschlecht

	A6.1.1	A6.1.2	A6.1.3	A6.1.4	A6.2.1	A6.2.2	A6.2.3	A6.2.4	A6.3.1	A6.3.2	A6.3.3	A6.3.4
p-Wert (gender)	0,982	0,969	0,231	0,823	0,353	0,769	0,833	0,424	0,814	0,911	0,844	0,063
p-Wert (mean, Nicht-Löser)	0,386	0,206	0,854	0,731	0,907	0,059	0,798	0,558	0,567	0,463	0,154	0,379

	A6.4.1	A6.4.2	A6.4.3	A6.4.4								
p-Wert (gender)	0,679	0,196	0,108	0,372								
p-Wert (mean, Nicht-Löser)	0,912	0,948	0,404	0,909								

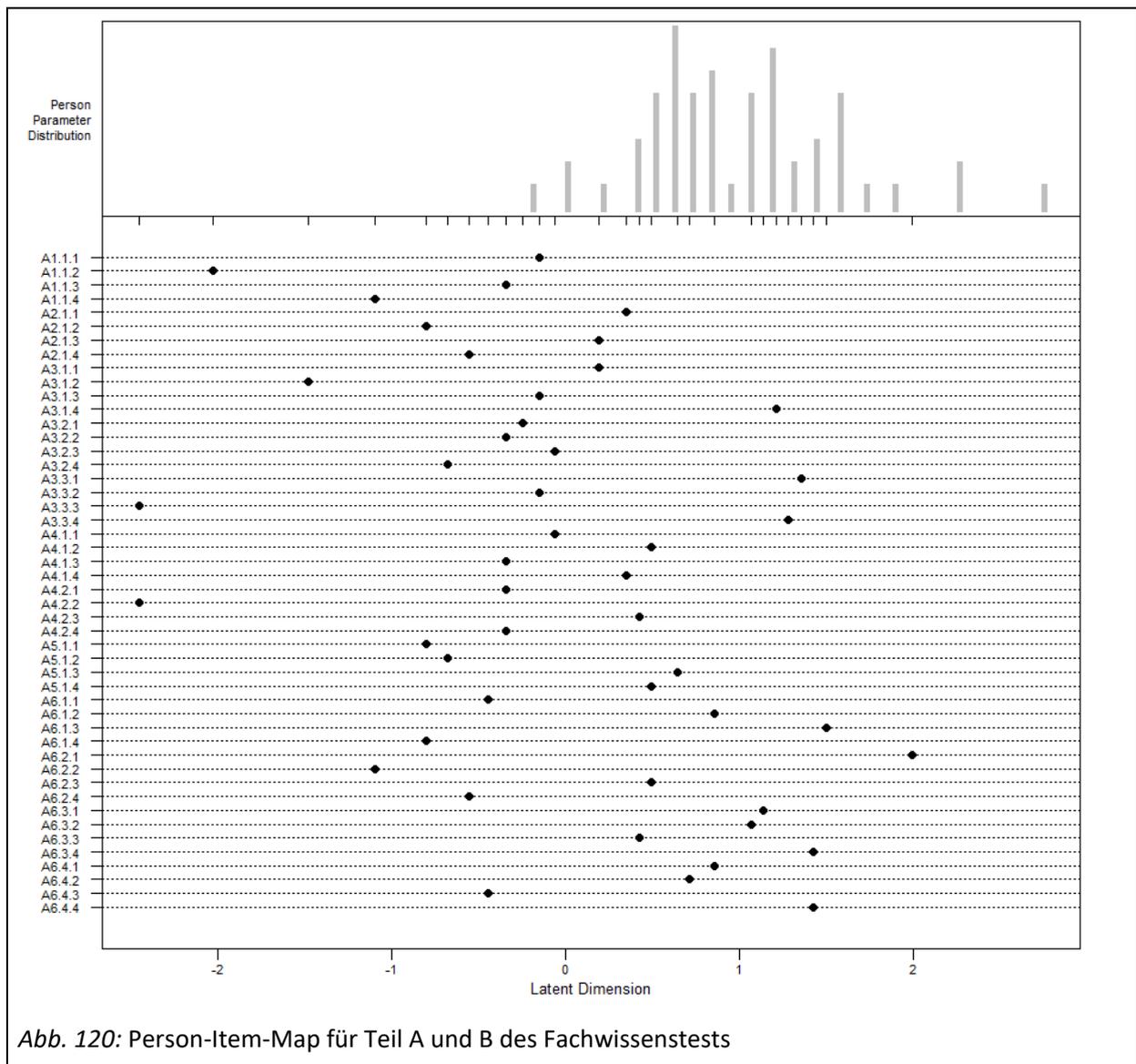


Abb. 120: Person-Item-Map für Teil A und B des Fachwissenstests

Tab. 62: Test auf Normalverteilung der Einzelteile des Beurteilungstests

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistik	df	Signifikanz	Statistik	df	Signifikanz
BT_A ORM	,206	59	,000	,903	59	,000
BT_B ORM	,154	59	,001	,945	59	,010
BT_C ORM	,190	59	,000	,946	59	,011
BT_D ORM	,169	59	,000	,935	59	,004
BT_E ORM	,131	59	,013	,935	59	,004

a. Signifikanzkorrektur nach Lilliefors

9 Literaturverzeichnis

- Abendroth-Smith, J. & Kras, J. (1999). More b-boat: The volleyball spike. *Journal of Physical Education, Recreation and Dance*, 70(3), 56–59.
- Adamczewski, H. (1995). Untersuchungsergebnisse vom Messplatz Speerwurf. *Die Lehre der Leichtathletik*, 17, 97–100.
- Adamczewski, H. & Perlt, B. (2003). Zu ausgewählten Problemen der Technik im Speerwurf von Spezialisten und Mehrkämpfern. *Zeitschrift für Angewandte Trainingswissenschaft*, 10(2), 44–74.
- Albanese, M. A., Kent, T. H. & Whitney, D. R. (1979). Cluing in multiple-choice test items with combinations of correct responses. *Academic Medicine*, 948–950.
- Annett, J. (1995). Motor imagery: Perception or action? *Neuropsychologia*, 33, 1395–1417.
- Artelt, C. & Gräsel, C. (2009). Diagnostische Kompetenz von Lehrkräften. *Zeitschrift für Pädagogische Psychologie*, 23(3–4), 157–160.
- Aschebrock, H. & Pack, R.-P. (2004). *Leisten und Leistung im Sportunterricht der Sekundarstufe I*. Aachen: Meyer & Meyer.
- Atienza, F., Balaguer, I. & Garcia-Merita, M. L. (1994). Factor analysis and reliability of the movement imagery questionnaire. *Perceptual and Motor Skills*, 78, 1323–1328.
- Bakker, M., de Lange, F., Stevens, J., Toni, L. & Bloem, B. (2007). Motor imagery of gait: A quantitative approach. *Experimental Brain Research*, 179(4), 497–504.
- Ballreich, R. (1981). Analyse und Ansteuerung sportmotorischer Techniken aus trainingsmethodischer und biomechanischer Sicht. *Leistungssport*, 6(11), 513–526.
- Ballreich, R. & Kuhlow, A. (Hrsg.). (1986). *Biomechanik der Leichtathletik. Biomechanik der Sportarten: Vol. 1*. Stuttgart: Enke.
- Balz, E. (2011). Zur Kompetenzorientierung im Sportunterricht. *Sportpädagogik*, 35(6), 52–56. Abgerufen am 24.09.2019 unter <http://www.schulsportsymposium.de/dl/Kompetenzorientierung%202.pdf>
- Banks, P., Moore, L., Liu, C. & Wu, B. (2004). Dynamic visual acuity: A review. *The South African Optometrist*, 63(2), 58–64.
- Barlett, R., Mueller, E., Lindinger, S., Brunner, F. & Morris, C. (1996). Three dimensional evaluation of the kinematic release parameters for javelin throwers of different skill level. *Journal of Applied Biomechanics*, 12(1), 58–71.
- Barrett, K. (1979). Observation for teaching and coaching. *Journal of Health, Physical Education and Recreation*, 50(1), 23–25.
- Barrett, K. (1984). A hypothetical model of observing as a teaching skill. *Journal of Teaching in Physical Education*, 3(1), 22–31.
- Bartolucci, F., Bacci, S. & Gnaldi, M. (2016). *Statistical analysis of questionnaires: A unified approach based on R and Stata*. Boca Raton: CRC Press.
- Bauersfeld, K.-H. & Schröter, G. (2016). *Grundlagen der Leichtathletik: Das Standardwerk für Ausbildung und Praxis* (6., aktualisierte Auflage). Aachen: Meyer & Meyer.
- Baumann, H. (1986). *Methoden der Fehleranalyse durch Bewegungsbeobachtung. Beiträge zur Bewegungsforschung im Sport: Vol. 9*. Bad Homburg: Limpert.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9(4), 469–520.

- Baumert, J. & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 29–54). Münster: Waxmann.
- Baumert, J., Stanat, P. & Demmrich, A. (2002). PISA 2000: Untersuchungsgegenstand, theoretische Grundlagen und Durchführung der Studie. In J. Baumert (Hrsg.), *PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 15–68). Opladen: Leske & Budrich.
- Ausgestaltung der inhaltlichen Prüfungsanforderungen für die Erste Staatsprüfung nach Kapitel II der Lehramtsprüfungsordnung I zu den einzelnen Fächern (Kerncurricula)https://www.km.bayern.de/download/13234_kerncurricula_fassung_mit_nd_151215.pdf (Bayerisches Staatsministerium für Unterricht und Kultus 15.10.2015).
- Beauchet, O., Annweiler, C., Assal, F., Bridenbaugh, S., Herrmann, F. R., Kressig, R. W. & Allali, G. (2010). Imagined Timed Up & Go test: A new tool to assess higher-level gait and balance disorders in older adults? *Journal of the Neurological Sciences*, 294(1–2), 102–106.
<https://doi.org/10.1016/j.jns.2010.03.021>
- Biscan, D. V. & Hofmann, S. J. (1976). Movement analysis as a generic ability of physical education teachers and students. *Research Quarterly*, 47(2), 161–163.
- Blischke, K., Schumacher, B. & Dausgs, R. (1993). Untersuchung zum Einfluss grafischer Orientierungshilfen auf die Fehleridentifikationsleistung von Experten und Novizen bei videografisch präsentierten sportmotorischen Bewegungsabläufen. In R. Dausgs (Hrsg.), *Schriften der Deutschen Vereinigung für Sportwissenschaft: Vol. 49. Aufmerksamkeit und Automatisierung in der Sportmotorik* (1. Auflage, S. 260–271). Sankt Augustin: Academia.
- Blischke, K. (1986). Zur Bedeutung bildhafter und verbaler Information für die Ausbildung einer Bewegungsvorstellung unter besonderer Berücksichtigung von Plazierungseffekten und Lernalter (Dissertation). Freie Universität, Berlin.
- Bond, T. G. & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (Third edition). New York: Routledge Taylor & Francis Group. Abgerufen am 24.09.2019 unter <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=1002030>
- Bortz, J. & Döring, N. (2009). *Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler* (6. Auflage). Berlin, Heidelberg: Springer.
- Bös, K. (Hrsg.). (2017a). *Handbuch Motorische Tests: Sportmotorische Tests, motorische Funktionstests, Fragebögen zur körperlich-sportlichen Aktivität und sportpsychologische Diagnoseverfahren* (3., überarbeitete und erweiterte Auflage). Göttingen: Hogrefe.
- Bös, K. (2017b). Testtheoretische Grundlagen. In K. Bös (Hrsg.), *Handbuch Motorische Tests: Sportmotorische Tests, motorische Funktionstests, Fragebögen zur körperlich-sportlichen Aktivität und sportpsychologische Diagnoseverfahren* (3. Auflage, S. 805–842). Göttingen: Hogrefe.
- Bös, K., Hänsel, F. & Schott, N. (2000). *Empirische Untersuchungen in der Sportwissenschaft: Planung, Auswertung, Statistik* (1. Auflage). Hamburg: Czwalina.
- Bös, K. & Wottawa, H. (1987). *Handbuch sportmotorischer Tests*. Göttingen: Hogrefe.
- Böttcher, J. & Kühl, L. (1996). Untersuchungen zum Speerwurf. *Die Lehre der Leichtathletik*, 35(31), 81–82.
- Böttcher, J. & Kühl, L. (1998). The technique of the best female javelin throwers in 1997. *New Studies in Athletics*, 13(1), 47–61.
- Bradshaw, E. J., Burnett, A. & Hume, P. A. (Hrsg.). (2012). *Scientific Proceedings of the 30th International Symposium on Biomechanics in Sports*.

- Bremer, D. & Sperle, N. (1984). Einführung in die Problemstellung der Fehleranalyse und Fehlerkorrektur. In D. Bremer & N. Sperle (Hrsg.), *Sport und Lernen: Vol. 9. Fehler, Mängel, Abweichungen im Sport: Von der fertigungs- zur handlungsorientierten Fehlerkorrektur und Mängelreduktion* (S. 6–13). Wuppertal: Putty.
- Brunner, M., Anders, Y., Hachfeld, A. & Krauss, S. (2011). Diagnostische Fähigkeiten von Mathematiklehrkräften. In M. Kunter, J. Baumert, W. Blum & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 215–234). Münster: Waxmann.
- Bühl, A. (2016). *SPSS 23: Einführung in die moderne Datenanalyse* (15., aktualisierte Auflage). Hallbergmoos: Pearson.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (3., aktualisierte und erw. Auflage). München: Pearson. Abgerufen am 24.09.2019 unter <http://lib.mylibrary.com/detail.asp?id=404890>
- Bühner, M. & Ziegler, M. (2009). *Statistik für Psychologen und Sozialwissenschaftler*. München: Pearson.
- Callow, N. & Hardy, L. (2004). The relationship between the use of kinaesthetic imagery and different visual imagery perspectives. *Journal of Sports Sciences*, 22(2), 167–177.
- Calmels, C. & Fournier, J. (2001). Duration of physical and mental execution of gymnastic routines. *The Sport Psychologist*, 15, 142–150.
- Campos, J., Brizuela, G. & Ramón, V. (2004). Three-dimensional kinematic analysis of elite javelin throwers at the 1999 IAAF World Championships in Athletics. *IAAF New Studies in Athletics*, 19(21), 47–57.
- Cerritelli, B., Maruff, P., Wilson, P. & Currie, J. (2000). The effect of an external load on the force and timing components of mentally represented actions. *Behavioural Brain Research*, 108, 91–96.
- Clark, J., Stamm, C. & Urquia, M. (1979). Development variability: The issue of reliability. In G. C. Roberts & K. M. Newell (Hrsg.), *Psychology of motor behavior and sport - 1978: Annual publication of the North American Society for the Psychology of Sport and Physical Activity ; papers in this volume were presented at the annual conference held May 21–24, 1978 at Florida State University* (S. 251–257). Champaign, Ill: Human Kinetics.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Erlbaum.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.
- Collet, C., Guillot, A., Lebon, F., MacIntyre, T. & Moran, A. (2011). Measuring motor imagery using psychometric, behavioral, and psychophysiological tools. *Exercise and Sport Sciences Review*, 39, 85–92.
- Connel, M., Sheridan, K. & Gardner, H. (2003). On abilities and domains. In R. J. Sternberg & E. L. Grigorenko (Hrsg.), *The psychology of abilities, competencies, and expertise* (S. 126–155). Cambridge, UK, New York: Cambridge University Press.
- Courtine, G., Papaxanthis, C., Gentili, R. & Pozzo, T. (2004). Gait-dependent motor memory facilitation in covert movement execution. *Cognitive Brain Research*, 22, 67–75.
- Cramer, C. (2012). *Entwicklung von Professionalität in der Lehrerbildung. Forschung*. Bad Heilbrunn, Tübingen: Klinkhardt.
- Daug, R., Blischke, K., Olivier, N. & Marschall, F. (1989). *Beiträge zum visuomotorischen Lernen im Sport. Schriftenreihe des Bundesinstituts für Sportwissenschaft: Vol. 65*. Schorndorf: Hofmann.
- Daug, R. (Hrsg.). (1984). *Visualisation sensomotorischer Lehrmedien. 1. Berliner Workshop 'Medien im Sport'*. Schriftenreihe der FVA des DSB. Berlin.

- Daug, R. & Blischke, K. (1984). Sensomotorisches Lernen. In K. Carl (Hrsg.), *Handbuch Sport: Wissenschaftliche Grundlagen von Unterricht und Training* (S. 381–420). Düsseldorf: Schwann.
- Daug, R., Blischke, K., Olivier, N. & Graf, M. (1983). *Zum Einfluß einer differenzierten Sollwertvorgabe auf das Blickverhalten von Trainern und Athleten bei der Fehleridentifikation*, BISp Köln.
- Dean, G. & M. P. (2003). The relationship between self-reports of imagery and spatial ability. *The British Journal of Psychology*, 94, 245–273.
- Decety, J. (1996). Do imagined and executed actions share the same neural substrate? *Cognitive Brain Research*, 3, 87–93.
- Decety, J. & Grézes, J. (1999). Neural mechanisms subserving the perception of human action. *Trends in Cognitive Science*, 3, 172–178.
- Decety, J., Jeannerod, M. & Prablanc, C. (1989). The timing of mentally represented actions. *Behavioral Brain Research*, 34(1–2), 35–45.
- Dunham, P. (1994). *Evaluation for physical education. Morton activity series*. Englewood, Colo: Morton.
- Eberspächer, H. (2001). *Mentales Training: Ein Handbuch für Trainer und Sportler* (5. Auflage). München: Copress.
- Ehrenstein, W. & Jendrusch, G. (2008). Dynamisches Sehen im Sport. *DOZ (Deutsche Optiker Zeitschrift) Optometrie & Fashion*, 63(5), 10–13.
- Eid, M., Gollwitzer, M. & Schmitt, M. (2017). *Statistik und Forschungsmethoden: Mit Online-Materialien* (5., korrigierte Auflage). Weinheim, Basel: Beltz.
- Eid, M. & Schmidt, K. (2014). *Testtheorie und Testkonstruktion*. Göttingen: Hogrefe.
- Elbing, E. (1978). *Pädagogische Handlungskompetenz: Entwurf eines Metakonzeptes theoriegeleiteter und wissenschaftsbezogener pädagogischer Praxis*. München: Ludwig-Maximilians-Universität.
- Engle, R. F. (2010). Wald, likelihood ratio, and lagrange multiplier tests in econometrics. In J. J. Heckman & E. Leamer (Hrsg.), *Handbooks in economics: Vol. 2. Handbook of econometrics* (S. 776–826). Amsterdam: Elsevier North Holland.
- Ennenbach, W. (1991). *Bild und Mitbewegung* (2. Auflage). *Betrifft Sonderband: Vol. 13*. Köln: bps.
- Erickson, G. (2007). *Sports vision: Vision care for the enhancement of sports performance*. St. Louis, Mo: Butterworth-Heinemann.
- Ericsson, K. A. & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (2nd edition). A *Bradford book*. Cambridge, Mass. [u.a.]: MIT Press.
- Ersland, L., Rosen, G., Lundervold, A., Smievoll, A. I., Tillung, T., Sundberg, H. & Ugdahl, K. (1996). Phantom limb imaginary fingertapping causes primary motor cortex activation: An fMRI study. *Neuro Report*, 8, 207–210.
- Eysel, U. (2007). Sehen und Augenbewegungen. In R. F. Schmidt & F. Lang (Hrsg.), *Physiologie des Menschen: Mit Pathophysiologie* (30. Auflage, S. 377–420). Berlin, Heidelberg: Springer.
- Fäh, B. & Straubhaar, F. (2011). *Checkliste Videoanalyse Disziplin Speerwurf*. St. Gallen. Abgerufen am 24.09.2019 unter Lehrgang Swiss Athletics website: http://www.speerschule.ch/docs/doc_videoanalyse3.pdf
- Farah, M. (1984). The neurological basis of mental imagery: A componential analysis. *Cognition*, 18, 245–272.
- Fetz, F. (1980). *Bewegungslehre der Leibesübungen* (2. Auflage). Bad Homburg: Limpert.
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47, 381–391.

- Frey, A. (2004). Die Kompetenzstruktur von Studierenden des Lehrerberufs: Eine internationale Studie. *Zeitschrift für Pädagogik*, 50(6), 903–925.
- Frey, A. (2006). Methoden und Instrumente zur Diagnose beruflicher Kompetenzen von Lehrkräften: Eine erste Standortbestimmung zu bereits publizierten Instrumenten. In C. Allemann-Ghionda & E. Terhart (Hrsg.), *Zeitschrift für Pädagogik. Beiheft: Vol. 51. Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern: Ausbildung und Beruf* (S. 30–46). Weinheim: Beltz.
- Fronsk, H. A. (2014). *Teaching cues for sport skills for secondary school students* (6th edition). Boston: Pearson.
- Fuchser, I. (2013). *Reihenbild Speerwurf - Zwölf Schlüsselpunkte für die Technik-Analyse*. Abgerufen am 24.09.2019 unter http://www.speerschule.ch/docs/doc_analysekeypoints.pdf
- Gabler, H. (2004). Kognitive Aspekte sportlicher Handlungen. In H. Gabler, J. R. Nitsch, R. Singer & J. Munzert (Hrsg.), *Sport und Sportunterricht: Vol. 2. Einführung in die Sportpsychologie* (4. Auflage, S. 34–63). Schorndorf: Hofmann.
- Gangstead, S. (1984, October). *A comparison of three methodological approaches to skill specific analytical training*, Jackson Hole, WY.
- Gangstead, S. & Beveridge, S. (1984). The implementation and evaluation of a methodological approach to qualitative sport skill analysis instruction. *Journal of Teaching in Physical Education*, 3(2), 60–70.
- Girardin, Y. & Hanson, D. (1967). Relationship between the ability to perform tumbling skills and the ability to diagnose performance errors. *Research Quarterly*, 38(4), 556–561.
- Girmes, R. (2006). Lehrerprofessionalität in einer demokratischen Gesellschaft: Über Kompetenzen und Standards in einer erziehungswissenschaftlich fundierten Lehrerbildung. In C. Allemann-Ghionda & E. Terhart (Hrsg.), *Zeitschrift für Pädagogik. Beiheft: Vol. 51. Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern: Ausbildung und Beruf* (S. 14–29). Weinheim: Beltz.
- Göhner, U. (1984). Lehren nach Funktionsphasen. Ein Rück- und Ausblick. In D. Kuhlmann (Hrsg.), *Kongreßberichte des Ausschusses Deutscher Leibeserzieher. Schüler im Sport, Sport für Schüler* (S. 96–106). Schorndorf: Hofmann.
- Gralla, V. (1996). Synchronoptisches Differenzierungs- und Identifizierungsvermögen im Sport: Ein methodischer Ansatz und erste Ergebnisse. In U. Bartmus, H. Heck, J. Mester, H. Schumann & G. Tidow (Hrsg.), *Aspekte der Sinnes- und Neurophysiologie im Sport: In memoriam Horst de Marées* (S. 305–313). Köln: Strauß.
- Gralla, V., Pfennig, H., Jendrusch, G., Tidow, G. & Heck, H. (1998). Zur Trainierbarkeit des synchronoptischen Identifizierungsvermögens. *Deutsche Zeitschrift für Sportmedizin*, 49(Sonderheft), 166–169.
- Grüsser, O.-J. & Grüsser-Cornhels, U. (1997). Gesichtssinn und Okulomotorik. In R. F. Schmidt & Schmidt-Thews (Hrsg.), *Physiologie des Menschen* (27. Auflage, S. 278–315). Berlin: Springer.
- Guillot, A. & Collet, C. (2005). Duration of mentally simulated movement: A review. *Journal of Motor Behavior*, 37(1), 10–20.
- Haase, W. & Bossow, B. (2004). Sehschärfe. In H. Kaufmann (Hrsg.), *Strabismus* (3. Auflage, S. 75–97). Stuttgart: Thieme.
- Haberkorn, C. & Plaß, R. (1992). *Leichtathletik: Sprung, Wurf, Stoß* (1. Auflage). *Spezielle Didaktik der Sportarten Bd. 2*. Frankfurt am Main: Diesterweg.
- Haladyna, T. M. (2004). *Developing and validating multiple-choice test items* (3rd edition). London, New York: Routledge Taylor & Francis Group.

- Hale, B. D. (1982). The effects of internal and external imagery on muscular and ocular concomitants. *Journal of Sport Psychology, 4*, 379–387.
- Hale, B. D. (1998). *Imagery training: A guide for sports people*. Leeds: National Coaching Foundation.
- Hale, B. D., Seiser, L., McGuire, E. J. & Weinrich, E. (2005). Mental imagery. In J. Taylor & G. S. Wilson (Hrsg.), *Applying sport psychology: Four perspectives* (S. 117–135). Champaign, IL: Human Kinetics.
- Hall, C., Mack, D., Paivio, A. & Hausenblas, H. (1998). Imagery use by athletes: Development of the Sport Imagery Questionnaire. *International Journal of Sport Psychology, 29*, 73–89.
- Hall, C., Pongrac, J. & Buckolz, E. (1985). The measurement of imagery ability. *Human Movement Science, 4*, 107–118.
- Hall, C., Rodgers, W. & Barr, K. (1990). The use of imagery by athletes in selected sports. *The Sport Psychologist, 4*(1–10).
- Hall, C. R. & Martin, K. A. (1997). Measuring movement imagery abilities: A revision of the Movement Imagery Questionnaire. *Journal of Mental Imagery, 21*(2), 143–154.
- Hall, C. R. & Pongrac, J. (1983). *Movement Imagery Questionnaire*. London, Ontario: University of Western Ontario.
- Hardy, L. (1997). The Coleman Roberts Griffith address: Three myths about applied consultancy work. *Journal of Applied Sport Psychology, 9*, 277–294.
- Hardy, L. & Callow, N. (1999). Efficacy of external and internal visual imagery perspectives for the enhancement of performance on tasks in which form is important. *Journal of Sport and Exercise Psychology, 21*, 95–112.
- Harnes, E. (1990a). Verwendung biomechanischer Meßdaten in der Trainingspraxis von Speerwerferinnen. *Die Lehre der Leichtathletik, 29*(48), 19–21.
- Harnes, E. (1990b). Verwendung biomechanischer Meßdaten in der Trainingspraxis von Speerwerferinnen. *Die Lehre der Leichtathletik, 29*(49), 19–22.
- Hartig, J. (2008). Kompetenzen als Ergebnisse von Bildungsprozessen. In N. Jude, J. Hartig & E. Klieme (Hrsg.), *Bildungsforschung: Vol. 26. Kompetenzerfassung in pädagogischen Handlungsfeldern: Theorien, Konzepte und Methoden* (S. 15–26). Bonn, Berlin.
- Hartig, J. & Jude, N. (2007). Empirische Erfassung von Kompetenzen und psychometrische Kompetenzmodelle. In J. Hartig & E. Klieme (Hrsg.), *Bildungsforschung: Vol. 20. Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik: Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung* (S. 17–36). Bonn, Berlin.
- Hartig, J. & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 127–143). Berlin, Heidelberg: Springer.
- Hartmann, C. (1999). Meines Merkmale der Bewegungskoordination als Kategorien zur Bewegungsbeobachtung und Bewegungsbeurteilung. In J. Krug & C. Hartmann (Hrsg.), *Sport und Wissenschaft: Vol. 8. Praxisorientierte Bewegungslehre als angewandte Sportmotorik: Symposiumsbericht Kurt-Meinel-Symposium, Universität Leipzig, Sportwissenschaftliche Fakultät, 1. bis 2. Dezember 1998* (1. Auflage, S. 106–116). Sankt Augustin: Academia.
- Hausenblas, H., Hall, C., Rodgers, W. & Monroe, K. (1999). Exercise imagery: Its nature and measurement. *Journal of Applied Sport Psychology, 11*, 171–180.
- Hay, J. G. & Reid, J. G. (1982). *The anatomical and mechanical bases of human motion*. Englewood Cliffs: Prentice Hall.
- Hayes, A. F. & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures, 1*(1), 77–89.

- Hecker, G. (1975). Probleme der Beobachtungsmethoden in Unterrichtsforschung. In H. Rieder, H. Eberspächer, K. Feige & E. Hahn (Hrsg.), *Internationales Symposium der Arbeitsgemeinschaft für Sportpsychologie in der Bundesrepublik Deutschland: Vol. 1. Empirische Methoden in der Sportpsychologie: Bericht über das 1. Internationale Symposium vom 9. bis 11. Oktober 1974 in Heidelberg* (1. Auflage, S. 79–82). Schorndorf: Hofmann.
- Helmke, A. (2014). *Unterrichtsqualität und Lehrerprofessionalität: Diagnose, Evaluation und Verbesserung des Unterrichts* (5. Auflage). *Unterricht verbessern - Schule entwickeln*. Seelze-Velber: Klett & Kallmeyer.
- Helmke, A., Hosenfeld, I. & Schrader, F.-W. (2004). Vergleichsarbeiten als Instrument zur Verbesserung der Diagnosekompetenz von Lehrkräften. In R. Arnold & C. Grieser (Hrsg.), *Schulleitung und Schulentwicklung: Voraussetzungen, Bedingungen, Erfahrungen* (S. 119–144). Baltmannsweiler: Schneider.
- Helsen, W. & Pauwels, J. M. (1993). A cognitive approach to visual search in sport. In D. Brogan, A. Gale & K. Carr (Hrsg.), *Visual Search 2* (S. 379–388). London: Taylor & Francis.
- Hofmann, S. J. & Armstrong, C. (1975). Effect of pretraining on performance error identification. In C. Bard, M. Fleury & J. Salmela (Hrsg.), *Mouvement: Actes du 7 symposium Canadien en apprentissage psychomoteur et psychologie du sport* (S. 209–214). Quebec: Revue de L'Association des Professionnels de L'Activité Physique du Quebec.
- Holmes, P. & Collins, D. (2001). The PETTLEP approach to motor imagery: A functional equivalence model for sport psychologists. *Journal of Applied Sport Psychology*, 13(1), 60–83.
- Höner, O. (2003). *Entscheidungshandeln im Sportspiel Fußball. Wissenschaftliche Schriftenreihe des Deutschen Sportbundes: Vol. 34*. Schorndorf, Bielefeld: Hofmann.
- Hotz, A. (1997). *Qualitatives Bewegungslernen: Bewegungsspielräume erleben, erkennen und gestalten* (3. Auflage). Bern: Schweizerischer Verband für Sport in der Schule.
- Hu, A. (2008). Kompetenzorientierung, Bildungsstandards und fremdsprachliches Lernen - Herausforderungen an die Fremdsprachenforschung. Abgerufen am 24.09.2019 unter http://www.dgff.de/fileadmin/user_upload/dokumente/Sonstiges/Kompetenzpapier_DGFF.pdf
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60, 581–592.
- Isaac, A., Marks, D. F. & Russell, D. G. (1986). An instrument for assessing imagery of movement: The Vividness of Movement Imagery Questionnaire (VMIQ). *Journal of Mental Imagery*, 10(4), 23–30.
- Janssen, J. & Laatz, W. (2017). *Statistische Datenanalyse mit SPSS: Eine anwendungsorientierte Einführung in das Basissystem und das Modul Exakte Tests* (9., überarbeitete und erweiterte Auflage). Berlin, Heidelberg: Springer Gabler.
- Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, 17(2), 187–202.
- Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *Neuro Image*, 14, 103–109.
- Jeannerod, M. & Decety, J. (1995). Mental motor imagery: A window into the representational stages of action. *Current Opinion in Neurobiology*, 5, 727–732.
- Jendrusch, G. (2006). Sportbezogene Leistungen des visuellen Systems. In U. Bartmus, G. Jendrusch, T. Henke & P. Platen (Hrsg.), *In Memoriam Horst de Marées anlässlich seines 70. Geburtstages: Beiträge aus Sportmedizin, Trainings- und Bewegungswissenschaft* (1. Auflage, S. 55–74). Köln: Strauß.
- Jendrusch, G. (2009). Sportspiele und visuelle Leistungsfähigkeit - Bochumer Perspektiven. In H.-F. Voigt (Hrsg.), *An der RUB - Sportpraxis nachgedacht. Sportspießforschung und -ausbildung in Bochum: Was war, was ist und was sein könnte* (S. 117–138). Hamburg: Czwalina.

- Kaminski, G. (1975). Einige Probleme der Beobachtung sportmotorischen Verhaltens. In H. Rieder, H. Eberspächer, K. Feige & E. Hahn (Hrsg.), *Internationales Symposium der Arbeitsgemeinschaft für Sportpsychologie in der Bundesrepublik Deutschland: Vol. 1. Empirische Methoden in der Sportpsychologie: Bericht über das 1. Internationale Symposium vom 9. bis 11. Oktober 1974 in Heidelberg* (1. Auflage, S. 43–65). Schorndorf: Hofmann.
- Kasess, C., Windischberger, C., Cunnington, R., Lanzenberger, R., Pezawas, L. & Moser, E. (2008). The suppressive influence of SMA on M1 in motor imagery revealed by fMRI and dynamic causal modeling. *Neuro Image*, 40, 828–837.
- Kelava, A. & Moosbrugger, H. (2012). Deskriptivstatistische Evaluation von Items (Itemanalyse) und Testwertverteilungen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (2. Auflage, S. 75–102). Berlin, Heidelberg: Springer.
- Killing, W. (2009). *Trainings- und Bewegungslehre des Hochsprungs* (2. Auflage). *Wissenschaftliche Berichte und Materialien / Bundesinstitut für Sportwissenschaft: Vol. 2004,09*. Köln: Strauß.
- Killing, W. (Hrsg.). (2011). *Jugendleichtathletik Wurf: Offizieller Rahmentrainingsplan des Deutschen Leichtathletik-Verbandes für die Wurfdisziplinen im Aufbautraining*. Münster: Philippka.
- Killing, W. (Hrsg.). (2015). *Jugendleichtathletik Mehrkampf: Rahmentrainingsplan des Deutschen Leichtathletik-Verbandes für die Mehrkampfdisziplinen im Aufbautraining*. Münster: Philippka.
- Klieme, E., Avenarius, H., Blum, W., Döbrich, P., Gruber, H., Prenzel, M. & Vollmer, H. J. (2007). *Zur Entwicklung nationaler Bildungsstandards*. Bonn, Berlin. Abgerufen am 24.09.2019 unter www.bmbf.d/pub/zur_entwicklung_nationaler_bildungsstandards.pdf
- Klieme, E. & Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen: Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik*, 52(6), 876–903.
- Klieme, E., Maag Merki, K. & Hartig, J. (2007). Kompetenzbegriff und Bedeutung von Kompetenzen im Bildungswesen. In J. Hartig & E. Klieme (Hrsg.), *Bildungsforschung: Vol. 20. Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik: Eine Expertise im Auftrag des Bundesministeriums für Bildung und Forschung* (S. 5–15). Bonn, Berlin.
- Klieme, E. & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcome-orientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54(2), 222–237.
- Knudson, D. V. (2013). *Qualitative analysis of human movement* (3rd edition). Leeds: Human Kinetics.
- Kohl, K. (1990). Bewegung sehen und beurteilen. *Sportpädagogik*, 14(1), 38–40.
- Koller, I., Alexandrowicz, R. & Hatzinger, R. (2012). *Das Rasch-Modell in der Praxis: Eine Einführung mit eRm*. Wien: Facultas.
- Koller, I., Maier, M. & Hatzinger, R. (2015). An empirical power analysis of quasi-exact tests for the Rasch model: Measurement invariance in small samples. *European Journal of Research Methods for the Behavioral and Social Sciences*, 11(2), 45–54.
- Krauss, S., Kunter, M., Brunner, M., Baumert, J., Blum, W., Neubrand, M. & Löwen, K. (2004). COACTIV: Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz. In J. Doll & M. Prenzel (Hrsg.), *Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung* (S. 31–53). Münster: Waxmann.
- Kubinger, K. D. (2011). Testorientierte Leistungsdiagnostik: Grundlagen und Probleme, Verfahrensinventar und besondere Einsatzgebiete. In L. F. Hornke (Hrsg.), *Enzyklopädie der Psychologie: Themenbereich B Methodologie und Methoden* (S. 1–50). Göttingen: Hogrefe.
- Kultusministerkonferenz (2004). Standards für die Lehrerbildung: Bildungswissenschaft, Kultusministerkonferenz 16.12.2004.

- Kultusministerkonferenz (2013). Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung, Kultusministerkonferenz 16.05.2013.
- Kunz, H. R. (2001). *Leichtathletik - Fehlerkorrektur*. Magglingen: BASPO.
- Langeheine, R., Pannekoek, J. & de Po, F. von. (1996). Bootstrapping goodness-of-fit measures in categorical data analysis. *Sociological Methods and Research*, 24, 492–516.
- Lehmann, F. (2005). Zum technikorientierten Krafttraining in den leichtathletischen Wurfdisziplinen am Beispiel des Speerwurfs. *Zeitschrift für Angewandte Trainingswissenschaft*, 12(1), 28–38.
- Lehmann, F. (2010). Biomechanical analysis of the javelin throw at the 2009 IAAF World Championships in Athletics. *New Studies in Athletics*, 25(3–4), 61–77.
- Lehmann-Grube, S. & Nickolaus, R. (2009). Professionalität als kognitive Disposition. In O. Zlatkin-Troitschanskaia, K. Beck, D. Sembill, R. Nickolaus & R. Mulder (Hrsg.), *Lehrprofessionalität: Bedingungen, Genese, Wirkungen und ihre Messung* (S. 59–70). Weinheim: Beltz.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Auflage). Weinheim: Beltz.
- Linacre, J. M. (2011). A user's guide to Winstep (R) Ministep: Rasch-model computer programs (3.72.3). Abgerufen am 24.09.2019 unter <http://ifile.hkedcity.net/1/001/950/public/Secondary/EI0020070012/winsteps.pdf>
- Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, 16, 878.
- Lindner, M. A., Strobel, B. & Köller, O. (2015). Multiple-Choice-Prüfungen an Hochschulen? *Zeitschrift für Pädagogische Psychologie*, 29(3–4), 133–149. <https://doi.org/10.1024/1010-0652/a000156>
- Logan, G. A. & McKinney, W. C. (1977). *Anatomic kinesiology* (2nd edition). Dubuque, Iowa: W.C. Brown Co.
- Lotze, M. & Halsband, U. (2006). Motor imagery. *Journal of Physiology*, 99, 386–395.
- Lotze, M., Montoya, P., Erb, M., Hülsmann, E., Flor, H., Klose, U. & Grodd, W. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: An fMRI study. *Journal of Cognitive Neuroscience*, 11, 491–501.
- Ludvig, E. & Miller, J. (1953). *A study of dynamic visual acuity*.
- Madan, C. & Singhal, A. (2012). Motor imagery and higher-level cognition: Four hurdles before research can sprint forward. *Cognitive Processing*, 13, 211–229.
- Mahoney, M. & Avenier, M. (1977). Psychology of the elite athlete: An exploratory study. *Cognitive Therapy and Research*, 2, 135–141.
- Mair, P., Hatzinger, Reinhold, Maier, Marko & Rusch, T. (2013). eRm - Extended Rasch modeling. R package v.0.15-7. Abgerufen am 24.09.2019 unter <https://cran.r-project.org/web/packages/eRm/>
- Maraun, H.-K. (1981). *Analysieren und Planen als Handlungsprobleme des Sportlehrers. Beiträge zur Lehre und Forschung im Sport: Vol. 82*. Schorndorf: Hofmann.
- Marées, H. d. & Mester, J. (1991). *Sportphysiologie* (2. Auflage). *Studienbücher Sport*. Frankfurt am Main: Diesterweg.
- Marschall, F., Blichke, K., Müller, H. & Reiser, M. (1996). Fehleridentifikation und Fehlerkorrektur beim sportmotorischen Lernen und Techniktraining. In R. Daus (Hrsg.), *Schriften der Deutschen Vereinigung für Sportwissenschaft: Vol. 73. Kognition und Motorik: Vom 19.–21.1.1995 in Saarbrücken* (S. 195–200). Hamburg: Czwalina.
- Martin, K. A., Moritz, S. & Hall, C. R. (1999). Imagery use in sport: A literature review and applied model. *The Sport Psychologist*, 13, 245–268.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149–174.

- Matschinger, H. (o.A.d.J.). *Entwicklung und die Überprüfung der psychometrischen Eigenschaften eines standardisierten Verfahrens zur Messung von Empowerment im Prozess der psychiatrischen Behandlung von Patienten mit affektiven (ICD-10 F30-F39) und schizophrenen (ICD-10 F20-F29) Störungen: Abschlussbericht*. Leipzig. Abgerufen am 24.09.2019 unter Universität Leipzig website: http://www.forschung-patientenorientierung.de/files/abschlussbericht__01gx0743_anhang_iii_methodenbeschreibung_irt_analysen.pdf-u1tC0qvFY82bmkA
- Mayer, J. & Hermann, H.-D. (2011). *Mentales Training: Grundlagen und Anwendung in Sport, Rehabilitation, Arbeit und Wirtschaft* (2. Auflage). Berlin, Heidelberg: Springer.
- McAvinue, L. & Robertson, I. (2008). Measuring motor imagery ability: A review. *European Journal of Cognitive Psychology*, 20(2), 232–251.
- McAvinue, L. & Robertson, I. (2009). Motor imagery: A multidimensional ability. *Journal of Mental Imagery*, 33(1–2), 99–120.
- McClelland, D. C. (1973). Testing for competence rather than for 'intelligence'. *American Psychologist*, 28, 1–14.
- Meinel, K. (1960). *Bewegungslehre: Versuch einer Theorie der sportlichen Bewegung unter pädagogischem Aspekt*. Berlin: Volk und Wissen.
- Meinel, K., Schnabel, G. & Krug, J. (Hrsg.). (2015). *Bewegungslehre - Sportmotorik: Abriss einer Theorie der sportlichen Motorik unter pädagogischem Aspekt* (12. Auflage). Aachen: Meyer & Meyer.
- Mero, A., Komi, P., Korjus, T., Navarro, E. & Gregor, R. J. (1994). Body segment contributions to javelin throwing during final thrust phases. *Journal of Applied Biomechanics*, 10(2), 166–177.
- Mester, J. (1988). *Diagnostik von Wahrnehmung und Koordination im Sport: Lernen von sportlichen Bewegungen*. Schorndorf: Hofmann.
- Mewes, N., Kellermann, M., Ehrenstein, W. & Jendrusch, G. (2008). Veränderungen der dynamischen Sehleistung bei körperlicher und psychischer Belastung. *Deutsche Zeitschrift für Sportmedizin*, 59(6), 141–145.
- Meyer, H. (2012). *Leitfaden Unterrichtsvorbereitung: [der neue Leitfaden]* (6. Auflage). Berlin: Cornelsen-Scriptor.
- Moreau, D., Clerc, J., Many-Dannay, A. & Guerrien, A. (2010). Assessing movement imagery ability: Self-report questionnaires vs. performance-based tests. *Europe's Journal of Psychology*, 6(4), 93–109.
- Morrison, C. (1994). Comparison of nationality, gender and type of instruction on the acquisition and retention of qualitative analysis of movement ability. In F. I. Bell & G. van Gyn (Hrsg.), *Access to active living: Proceedings for the 10th Commonwealth & International Scientific Congress, 10–14 August 1994, University of Victoria, Victoria, British Columbia, Canada = Accès à une vie active* (S. 169–174). [Victoria, B.C.]: The School.
- Morrison, C. (2000). Why don't you analyze the way I analyze?. *Journal of Physical Education, Recreation and Dance*, 71(1), 22–25.
- Morrison, C. & Reeve, J. (1989). Effect of different instructional videotape units on undergraduate physical education majors' qualitative analysis of skill. *Perceptual and Motor Skills*, 69, 111–114.
- Morrison, C. & Reeve, J. (1992). Perceptual style and instruction in the acquisition of qualitative analysis of movement by majors in elementary education. *Perceptual and Motor Skills*, 74(2), 579–583.
- Morrison, C. & Reeve, J. (1993). A framework for writing and evaluating critical performance cues in instructional materials for physical education. *Physical Educator*, 50(3), 132–135.

- Morriss, C. & Barlett, R. (1996). Biomechanical factors critical for performance in the men's Javelin throw. *Sports Medicine*, 21(6), 438–446.
- Muenchen, R. A. (2011). *R for SAS and SPSS Users* (2nd edition). *Statistics and computing*. New York, NY: Springer.
- Munroe, K., Giacobbi, P., Hall, C. & Weinberg, R. (2000). The four Ws of imagery use: Where, when, why and what. *The Sport Psychologist*, 14, 119–137.
- Munzert, J. (1992). Motorik-Repräsentation, Bewegungswissen und Bewegungshandeln. *Sportwissenschaft*, 22, 344–356.
- Munzert, J. (2001a). Bewegungsvorstellungen – Bewegungshandlungsvorstellungen. In D. Hackford (Hrsg.), *Handlungspsychologische Forschung für die Theorie und Praxis der Sportpsychologie* (S. 49–63). Köln: bps.
- Munzert, J., Lorey, B. & Zentgraf, K. (2009). Cognitive motor processes: The role of motor imagery in the study of motor representations. *Brain Research Reviews*, 60(2), 306–326.
- Munzert, J. (2001b). Vorstellung und Bewegung. In J. R. Nitsch & H. Allmer (Hrsg.), *Denken - Sprechen - Bewegen: Bericht über die 32. Tagung der Arbeitsgemeinschaft für Sportpsychologie (asp) vom 1. bis 3. Juni 2000 in Köln* (1. Auflage, S. 41–56). Köln: bps.
- Munzert, J. (2008). Does level of expertise influence imagined durations in open skills? Played versus imagined durations of badminton sequences. *International Journal of Sport and Exercise Psychology*, 6, 24–38.
- Munzert, J., Dültgen, K. & Möllmann, H. (2000). Individuelle Merkmale von Bewegungsvorstellungen: Eine explorative Untersuchung im Badminton. *Psychologie und Sport*. (7), 15–25.
- Munzert, J. & Reiser, M. (2003). Vorstellung und mentales Training. In H. Mechling & J. Munzert (Hrsg.), *Handbuch Bewegungswissenschaft - Bewegungslehre* (S. 219–230). Schorndorf: Hofmann.
- Murakami, M., Tanabe, S., Ishikawa, M., Isolehto, J., Komi P. & Ito, A. (2006). Biomechanical analysis of the javelin at the 2005 IAAF World Championships in Athletics. *IAAF New Studies in Athletics*, 21(2), 67–80.
- Murphy, K. R. & Davidshofer, C. O. (1998). *Psychological testing: Principles and applications*. Upper Saddle River, NJ: Prentice-Hall.
- Naito, E. (1994). Controllability of motor imagery and transformation of visual imagery. *Perception Motor Skills*, 78, 479–487.
- Nelson, J. K., Thomas, J. R., Nelson, K. R. & Abraham, P. C. (1986). Gender differences in children's throwing performance: Biology and environment. *Research Quarterly for Exercise and Sport*, 57(4), 280–287. <https://doi.org/10.1080/02701367.1986.10608088>
- Neuendorf, K. A. (2010). *The content analysis guidebook* (9th edition). Thousand Oaks: Sage.
- Neumaier, A. (1979). Visuelle Informationsaufnahme sportlicher Bewegungsabläufe: Eine Untersuchung zum Blickverhalten beim Betrachten von filmisch dargebotenen Übungsteilen aus dem Bodenturnen (Dissertation), Köln.
- Neumaier, A. (1988). *Bewegungsbeobachtung und Bewegungsbeurteilung im Sport* (1. Auflage). *Schriften der Deutschen Sporthochschule Köln: Vol. 21*. Sankt Augustin: Academia.
- Neumaier, A. & Jendrusch, G. (1999). Aktuelle Positionen zum Bewegungssehen im Sport. In J. Krug & C. Hartmann (Hrsg.), *Sport und Wissenschaft: Vol. 8. Praxisorientierte Bewegungslehre als angewandte Sportmotorik: Symposiumsbericht Kurt-Meinell-Symposium, Universität Leipzig, Sportwissenschaftliche Fakultät, 1. bis 2. Dezember 1998* (1. Auflage, S. 128–141). Sankt Augustin: Academia.

- Neumann, K. (2014). Rasch-Analyse naturwissenschaftsbezogener Leistungstests. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 355–370). Berlin: Springer.
- Nishida, T., Katube, A. & Inomata, K. (1986). A new test for controllability of motor imagery: The examination of its validity and reliability. *Japanese Journal of Physical Education*, 31, 13–22.
- Orliaguet, J. & Coello, Y. (1998). Differences between actual and imagined putting movements in golf: A chronometric analysis. *International Journal of Sport Psychology*, 29, 157–169.
- Osborne, M. & Gordon, M. (1972). An investigation of the accuracy of ratings of a gross motor skill. *Research Quarterly*, 43(1), 55–61.
- Oser, F. (2001). Standards: Kompetenzen von Lehrpersonen. In F. Oser & J. Oelkers (Hrsg.), *Nationales Forschungsprogramm 33 - Wirksamkeit unserer Bildungssysteme. Die Wirksamkeit der Lehrerbildungssysteme: Von der Allrounderbildung zur Ausbildung professioneller Standards* (S. 215–342). Chur: Rüegger.
- Oser, F. & Renold, U. (2005). Kompetenzen von Lehrpersonen: Über das Auffinden von Standards und ihre Messung. In I. Gogolin, H.-H. Krüger, D. Lenzen & T. Rauschenbach (Hrsg.), *Zeitschrift für Erziehungswissenschaft Beiheft: Vol. 4. Standards und Standardisierungen in der Erziehungswissenschaft* (S. 119–140). Wiesbaden: VS Verlag.
- Parsons, L. M. (1987). Imagined spatial transformation of one's body. *Journal of Experimental Psychology: General*, 116(2), 172–191.
- Parsons, L. M. (1994). Temporal and kinematic properties of motor behavior reflected in mentally simulated action. *Journal of Experimental Psychology: Human Perception and Performance*, 20(4), 709–730.
- Petrakis, E. (1993). Analysis of visual search patterns of tennis teachers. In G. D'Ydewalle (Hrsg.), *Studies in visual information processing: Vol. 4. Perception and cognition: Advances in eye movement research* (S. 159–168). Amsterdam: North-Holland.
- Pinheiro, V. & Cai, S. (1999). Preservice teachers diagnosing live motor performance. *Research Quarterly for Exercise and Sport*, 70(1), A98–A99.
- Pöhlmann, R. (1986). *Motorisches Lernen: Psychomotorische Grundlagen der Handlungsregulation sowie Lernprozeßgestaltung im Sport* (1. Auflage). *Sportwissenschaft für die Praxis: Vol. 1*. Berlin: Sportverlag.
- Ponocny, I. (2001). Nonparametric Goodness-of-fit tests for the Rasch Model. *Psychometrika*, 66(3), 437–460.
- Radford, K. W. (1990). Movement education in physical education: A definitional effort. *Journal of Teaching in Physical Education*, 9(1), 1–24.
- Reed, C. (2002). Chronometric comparisons of imagery to action: Visualizing versus physically performing springboard dives. *Memory and Cognition*, 30(8), 1169–1178.
- Rentzsch, K. & Schütz, A. (2009). *Psychologische Diagnostik* (1. Auflage). Stuttgart: Kohlhammer.
- Rizopoulos, D. (2006). ltm: An R package for latent variable modeling and item response theory analyses. *Journal of Statistical Software*, 17(5), 1–25.
- Roach, N. T. & Lieberman, D. E. (2014). Upper body contributions to power generation during rapid, overhand throwing in humans. *The Journal of Experimental Biology*, 217(Pt 12), 2139–2149. <https://doi.org/10.1242/jeb.103275>
- Roach, N. T. & Richmond, B. G. (2015). Clavicle length, throwing performance and the reconstruction of the Homo erectus shoulder. *Journal of Human Evolution*, 80, 107–113. <https://doi.org/10.1016/j.jhevol.2014.09.004>

- Roach, N. T., Venkadesan, M., Rainbow, M. J. & Lieberman, D. E. (2013). Elastic energy storage in the shoulder and the evolution of high-speed throwing in Homo. *Nature*, 498(7455), 483–486. <https://doi.org/10.1038/nature12267>
- Roberts, R., Callow, N., Hardy, L., Markland, D. & Bringer, J. (2008). Movement imagery ability: Development and assessment of a revised version of the Vividness of Movement Imagery Questionnaire. *Journal of Sport and Exercise Psychology*, 30, 200–221.
- Rockmann, U. & Bömermann, H. (2006). *Grundlagen der sportwissenschaftlichen Forschungsmethoden und Statistik. Grundlagen der Sportwissenschaft: Vol. 2*. Schorndorf: Hofmann.
- Roland, P., Larsen, B., Lassen, N. & Skinhoj, E. (1980). Supplementary motor area and other cortical areas in organization of voluntary movements in man. *Journal of Neurophysiology*, 43, 118–136.
- Römer, J. (2012). Das Geheimnis des weiblichen Kurzwurfs. Abgerufen am 24.09.2019 unter <http://www.spiegel.de/gesundheit/ernaehrung/weitwurf-warum-werfen-maedchen-schlechter-als-jungen-a-855798.html>
- Rost, J. (1990). Rasch models in latent classes: An integration of two approaches to item analysis. *Applied Psychological Measurement*, 14(3), 271–282.
- Rost, J. (2004). *Lehrbuch Testtheorie - Testkonstruktion* (2. Auflage). *Psychologie Lehrbuch*. Bern: Huber.
- Rostock, J. (2003). *Haltung und Bewegung im Sport: Theorie, Aneignung und Diagnostik aus biomechanischer, bewegungs- und trainingswissenschaftlicher Sicht*. Chemnitz: TU-Chemnitz.
- Rostock, J. & Kirchner, G. (1999). Standpunkte zur kritischen Reflexion der Bewegungsmerkmale Meinels in der Literatur sowie zu ihrem Einsatz im Rahmen der Bewegungsbeobachtung. In J. Krug & C. Hartmann (Hrsg.), *Sport und Wissenschaft: Vol. 8. Praxisorientierte Bewegungslehre als angewandte Sportmotorik: Symposiumsbericht Kurt-Meinell-Symposium, Universität Leipzig, Sportwissenschaftliche Fakultät, 1. bis 2. Dezember 1998* (1. Auflage, S. 117–127). Sankt Augustin: Academia.
- Rushal, B. & Lippman, L. (1998). The role of imagery in physical performance. *International Journal of Sport Psychology*, 29, 57–72.
- Sage, G. (1984). *Motor learning and control: A neurophysiological approach*. Dubuque, Iowa: W.C. Brown.
- Schecker, H. (2014). Überprüfung der Konsistenz von Itemgruppen mit Cronbachs alpha. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung*. Berlin, Heidelberg: Springer. Abgerufen am 24.09.2019 unter <http://static.springer.com/sgw/documents/1426184/application/pdf/Cronbach>
- Schellig, D. & Hättig, H. A. (1993). Die Bestimmung der visuellen Merkspanne mit dem Block-Board. *Zeitschrift für Neuropsychologie*, 4(2), 104–112.
- Scheuneman, J. D. & Subhiyah, R. G. (1998). Evidence for the validity of a Rasch model technique for identifying differential item functioning. *Journal of Outcome Measurement*, 2(1), 33–42.
- Schmiemann, P. & Lücken, M. (2014). Validität - Misst mein Test, was er soll? In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 107–120). Berlin: Springer.
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, 8(4), 350–353.
- Schnell, D. (2010). Visuelles System. In H.-H. Dickhuth & G. Badtke (Hrsg.), *Sportmedizin für Ärzte: Lehrbuch auf der Grundlage des Weiterbildungssystems der Deutschen Gesellschaft für Sportmedizin und Prävention (DGSP)* (2. Auflage, S. 209–229). Köln: Deutscher Ärzte-Verlag.

- Schnitzler, A., Salenius, S., Salmelin, R., Jousmäki, V. & Hari, R. (1997). Involvement of primary motor cortex in motor imagery: A neuromagnetic study. *Neuro Image*, 3(6), 201–208.
- Schöllhorn, W., Beckmann, H., Janssen, D. & Michelbrink, M. (2009). Differenzielles Lehren und Lernen im Sport. *Sportunterricht*, 58(2), 36–40.
- Scholze-Stubenrecht, W. (2013). *Duden - die deutsche Rechtschreibung: [das umfassende Standardwerk] auf der Grundlage der aktuellen amtlichen Rechtschreibregeln* (26. Auflage). *Der Duden: Vol. 1*. Berlin: Dudenverlag.
- Schöpe, H. G. (1997). *Die Entwicklung der Bewegungsvorstellung im Gerätturnen: Eine sich ändernde Beziehung von Theorie und Praxis* (1. Auflage). *Sportwissenschaft und Sportpraxis: Vol. 107*. Hamburg: Czwalina.
- Schott, N. (2013). Test zur Kontrollierbarkeit der Bewegungsvorstellungsfähigkeit (TKBV) bei älteren Erwachsenen. *Zeitschrift für Gerontologie und Geriatrie*, 46(7), 663–672.
- Schott, N. (2012). Age-related differences in motor imagery: Working memory as a mediator. *Experimental Aging Research*, 38(5), 559–583. <https://doi.org/10.1080/0361073X.2012.726045>
- Schrader, F.-W. (2010). Diagnostische Kompetenz von Eltern und Lehrern. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4. Auflage, S. 95–100). Weinheim: Beltz.
- Schubert, F. & Zehl, U.-C. (1984). Möglichkeiten der Blickbewegungsregistrierung im Sport. *Medizin und Sport*, 24.
- Sharma, N., Pomeroy, V. & Baron, J. (2006). Motor imagery: A backdoor to the motor system after stroke? *Stroke*, 37(7), 1941–1952.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–22.
- Staatsinstitut für Schulqualität und Bildungsforschung München (ISB). Lehrplan Plus Gymnasium: Fachprofil Sport. Abgerufen am 24.09.2019 unter <https://www.lehrplanplus.bayern.de/fachprofil/gymnasium/sport>
- Stahl, E. (1988). Häufige Fehler und ihre Korrektur. Abgerufen am 24.09.2019 unter http://www.speerschule.ch/docs/doc_fehlerkorrektur-sjw.pdf
- Strobl, C. (2012). *Das Rasch-Modell: Eine verständliche Einführung für Studium und Praxis* (2. Auflage). *Sozialwissenschaftliche Forschungsmethoden*. Mering: Hampp.
- Strobl, C., Kopf, J. & Zeileis, A. (2015). A new method for detecting differential item functioning in the Rasch model. *Psychometrika*, 80(2), 289–316.
- Strüder, H. K., Jonath, U. & Scholz, K. (2016). *Leichtathletik: Trainings- und Bewegungswissenschaft - Theorie und Praxis aller Disziplinen* (2. korr. Auflage). Hellenthal: Strauß.
- Swert, K. d. (2012). Calculating inter-coder reliability in media content analysis using Krippendorff's Alpha. Abgerufen am 24.09.2019 unter <http://www.polcomm.org/wp-content/uploads/ICR01022012.pdf>
- Terhart, E. (2002). Standards für die Lehrerbildung: Eine Expertise für die Kultusministerkonferenz. ZKL-Texte Nr. 23. Abgerufen am 24.09.2019 unter http://miami.uni-muenster.de/servlets/DerivateServlet/Derivate-1151/Standards_fuer_die_Lehrerbildung_Eine_Expertise_fuer_die_Kultusministerkonferenz.pdf
- Terhart, E. (2007). Erfassung und Beurteilung beruflicher Kompetenzen von Lehrkräften. In M. Lüders & J. Wissinger (Hrsg.), *Forschung zur Lehrerbildung: Kompetenzentwicklung und Programmevaluation* (S. 38–62). Münster: Waxmann.

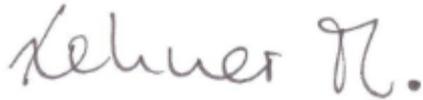
- Thomas, J. R., Alderson, J. A., Thomas, K. T., Campbell, A. C. & Elliott, B. C. (2010). Developmental gender differences for overhand throwing in Aboriginal Australian children. *Research Quarterly for Exercise and Sport*, 81(4), 432–441. <https://doi.org/10.1080/02701367.2010.10599704>
- Thomas, J. R. & French, K. E. (1985). Gender differences across age in motor performance: A meta-analysis. *Psychological Bulletin*, 98(2), 260–282.
- Tidow, G. (1981). Modell zur Technikschiilung und Bewegungsbeurteilung in der Leichtathletik: Technikbilder - Phasenstrukturen - Kriterienkataloge. *Leistungssport*, 11(4), 264–277.
- Tidow, G. (1983). *Beobachtung und Beurteilung azyklischer Bewegungsabläufe: Dargestellt am Beispiel der Disziplin Kugelstoßen* (1. Auflage). *Sportwissenschaft und Sportpraxis: Vol. 50*. Ahrensburg bei Hamburg, Dortmund: Czwalina.
- Tidow, G. (1993). Bewegungssehen im Sport - Möglichkeiten und Grenzen. In H.-F. Voigt (Hrsg.), *An der RUB - Sportpraxis nachgedacht. Bewegungen lesen und antworten* (1. Auflage, S. 15–72). Ahrensburg bei Hamburg: Czwalina.
- Tidow, G. (1996a). Model technique analysis sheets - Part X: The javelin throw. *New Studies in Athletics*, 11(1), 45–62.
- Tidow, G. (1996b). Zur Optimierung des Bewegungssehens im Sport. In U. Bartmus, H. Heck, J. Messter, H. Schumann & G. Tidow (Hrsg.), *Aspekte der Sinnes- und Neurophysiologie im Sport: In memoriam Horst de Marées* (S. 241–286). Köln: Strauß.
- Tidow, G., Koch, E. & Marées, H. d. (1984). Visuelle Probleme bei der Beurteilung azyklischer Bewegungsabläufe (Kugelstoßen). In D. Jeschke (Hrsg.), *Stellenwert der Sportmedizin in Medizin und Sportwissenschaft/Position of Sports Medicine in Medicine and Sports Science: 2. Symposium der Sektion „Sportmedizinische Forschung und Lehre an den Hochschulen“ des Deutschen Sportärztebundes (Deutsche Gesellschaft für Sportmedizin e.V.) Tübingen, 5.–8. Mai 1983* (S. 128–136). Berlin, Heidelberg: Springer.
- Trepel, M. (2004). *Neuroanatomie: Struktur und Funktion*. München u.a.: Urban & Fischer.
- Ulich, E. (1967). Über verschiedene Methoden des Lernens sensumotorischer Fertigkeiten. *Arbeitswissenschaft*, 6(2), 48–50.
- Ungerleider, S. & Golding, J. (1991). Mental practice among Olympic athletes. *Perceptual and Motor Skills*, 72, 1007–1017.
- Vandenberg, S. & Kuse, A. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills*, 47, 599–604.
- Verhelst, N. (2008). An efficient MCMC algorithm to sample binary matrices with fixed marginals. *Psychometrika*, 73(4), 708–728.
- Vonken, M. (2005). *Handlung und Kompetenz* (1. Auflage). Wiesbaden, Erfurt: VS Verlag.
- Wang, W.-C. & Chen, C.-T. (2005). Item parameter recovery, standard error estimates, and fit statistics of the Winsteps program for the family of Rasch models. *Educational and Psychological Measurement*, 65(3), 376–404.
- Weiler, A. & Lindemann, U. (2010). Werfen mit dem Speer. *Betrifft Sport*, 32(5), o.A.d.S.
- Weinberg, R. (1988). *The mental advantage: Developing your psychological skills in tennis*. Champaign, IL: Human Kinetics.
- Weineck, J. (2004). *Sportbiologie* (9. Auflage). Balingen: Spitta.
- Weinert, F. E. (1999). *Concepts of competence (Contribution within the OECD project Definition and selection of competencies: Theoretical and conceptual foundations (DeSeCo))*. Neuchatel: DeSeCo.

- Weinert, F. E. (2001a). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Hrsg.), *Defining and selecting key competencies: Theoretical and Conceptual Foundations* (S. 45–66). Seattle, Toronto: Hogrefe & Huber.
- Weinert, F. E. (2001b). Vergleichende Leistungsmessung in Schulen - Eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (2. Auflage, S. 17–31). Weinheim: Beltz.
- Weinert, F. E., Schrader, F.-W. & Helmke, A. (1990). Unterrichtsexpertise. In L.-M. Alisch, J. Baumert & K. Beck (Hrsg.), *Braunschweiger Studien zur Erziehungs- und Sozialarbeitswissenschaft: Vol. 28. Professionswissen und Professionalisierung* (S. 173–206). Braunschweig: Technische Universität.
- Westers, P. & Kelderman, H. (1991). Examining differential item functioning due to item difficulty and alternative attractiveness. *Psychometrika*, 57(1), 107–118.
- Wiemann, K. (1989). Zum Phänomen der Bewegungsvorstellung aus hirnbioologischer Sicht. In R. Daus, K.-H. Leist & H.-V. Ulmer (Hrsg.), *Symposium Ansätze interdisziplinärer Forschung im Bereich der Sportwissenschaft: Vol. 7. Motorikforschung aktuell* (S. 196–203). Clausthal-Zellerfeld: dvs.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwa, NJ: Erlbaum.
- Wollersheim, H.-W. (1993). *Kompetenzerziehung: Befähigung zur Bewältigung*. Frankfurt am Main, Berlin: Lang.
- Wolters, P. (2002). *Bewegungskorrektur im Sportunterricht* (2. Auflage). *Forum Sportwissenschaft: Vol. 2*. Schorndorf, Hamburg: Hofmann.

Versicherung (gem. §4 Abs. 3 Satz 1 Nr. 5 PromO):

Ich versichere hiermit

- an Eides statt, dass ich die Dissertation selbständig angefertigt, außer den im Schriftenverzeichnis sowie den Anmerkungen genannten Hilfsmitteln keine weiteren benutzt und die Herkunft der Stellen, die wörtlich oder sinngemäß aus anderen Werken übernommen sind, bezeichnet habe,
- dass ich die Dissertation nicht bereits in derselben oder einer ähnlichen Fassung an einer anderen Fakultät oder einer anderen Hochschule zur Erlangung eines akademischen Grades eingereicht habe.



Passau, den 26.09.2018