

Die Struktur epistemologischer Überzeugungen

Empirische und theoretische Analysen

Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Philosophie
der Philosophischen Fakultät III
der Universität des Saarlandes

vorgelegt von

Eric Klopp

aus Saarbrücken

Saarbrücken, 2014

Dekan:
Prof. Dr. Roland Brünken

Berichterstatter;
Prof. Dr. Robin Stark
Prof. Dr. Frank Schwab

Tag der Disputation: 20. Mai 2014

Inhaltsverzeichnis

1. Überblick zur Forschung	1
1.1 Einführung	1
1.2 Definition des Begriffs „epistemologische Überzeugungen“ und geschichtlicher Abriss des Konzepts	2
1.3 Persönlichkeitspsychologische Modelle epistemologischer Überzeugungen	7
1.3.1 Schommers Epistemological-Belief-Modell.....	7
1.3.2 Epistemologische Überzeugungen als Persönliche Theorien.....	18
1.3.3 Epistemologische Überzeugungen nach Moschner und Gruber	30
1.4 Epistemologische Ressourcen	38
1.5 Alternative Ansätze epistemologischer Überzeugungen	45
1.5.1 Psychologische Epistemologie von Royce.....	46
1.5.2 Theory of Integrated Domains (TIDE) – Framework	50
1.5.3 Epistemologische Überzeugungen und Metakognition.....	55
1.6 Pädagogische Relevanz epistemologischer Überzeugungen	60
1.6.1 Zusammenhang epistemologischer Überzeugungen und Lernstrategien.....	60
1.6.2 LIST-Fragebogen zur Erfassung von Lernstrategien im Studium	71
1.7 Epistemologische Überzeugungen als latente Variablen	75
2. Forschungsfragen	85
3. Methoden	88
3.1 Explorative Faktorenanalyse	88
3.2 Konfirmatorische Faktorenanalyse	95
3.2.1 Grundlegende CFA-Modelle.....	95
3.2.2 Bestimmung der Modellpassung.....	103
3.2.3 Modifikationen des Modells.....	107
3.2.4 Kreuzvalidierung der CFA-Modelle	112
3.3 Zusammenfassung der Auswertungsstrategien faktorenanalytischer Methoden 113	
3.4 Psychometrische Eigenschaften	115
3.4.1 Definition der Reliabilität.....	115
3.4.2 Voraussetzungen der Reliabilitätsschätzung.....	116
3.4.3 Schätzungen der internen Konsistenz	119
3.5 Auswertungsstrategie für psychometrische Fragestellungen	121
3.6 Validität	122
3.6.1 Definition und Bedeutung des Validitätsbegriffs.....	122
3.6.2 Konstruktvalidität auf Itemebene	123
3.6.3 Konstruktvalidität auf Ebene des Gesamttests	124
3.7 Auswertungsstrategie zur Validierung des EPI-Fragebogens	133
4. Ergebnisse zum EBQ	134
4.1 Faktorenanalytische Auswertung des EBQ	134
4.1.1 Vorbereitung der Daten.....	134
4.1.2 Auswertung nach dem Modell von Schommer.....	138
4.1.3 Explorative Faktorenanalyse des Schommer Fragebogens auf Itemebene	146
4.1.4 Psychometrische Eigenschaften und inhaltliche Interpretation der Faktoren	158
4.1.4.1 Skala F1: Autoritäres-einfaches Wissen	158
4.1.4.2 Skala F2: Existenz von Wahrheit	160
4.1.4.3 Skala F3: Integratives Wissen	161
4.1.4.4 Faktor 4: Lernmethodik.....	163
4.1.4.5 Faktor 5: Unsicherheit des Wissens	164

4.2 Diskussion der Analysen des EBQ	167
4.2.1 Das Modell von Schommer	167
4.2.2 Auswertung des EBQ auf Itemebene	169
4.2.3 Abgeleitete Skalen und deren psychometrische Eigenschaften	172
5. Ergebnisse zum EPI	182
5.1 Faktorenanalytische Auswertung des EPI.....	182
5.1.1 Vorbereitung der Daten	182
5.1.2 CFA des Gesamtmodells mit sieben Faktoren	188
5.1.3 Prüfung der einzelnen Faktoren	188
5.1.3.1 Faktor Reflexive Natur des Wissens	188
5.1.3.2 Faktor Absolutes Wissen.....	191
5.1.3.3 Faktor Soziale Komponente des Wissens	199
5.1.3.4 Faktor Value of Knowledge	201
5.1.3.5 Faktor Geschlechtsspezifische Wissenszugänge.....	203
5.1.3.6 Faktor Kulturspezifische Wissenszugänge.....	206
5.1.3.7 Faktor Lernen lernen	208
5.1.4 Beurteilung des revidierten Gesamtmodells	211
5.2 Diskussion der Analysen des EPI.....	219
5.3 Konstruktvalidierung des EPI	232
5.3.1 Faktorenstruktur des LIST	232
5.3.1.1 Vorbereitung der Daten	232
5.3.1.2 Konfirmatorische Faktorenanalysen des LIST.....	235
5.3.1.2.1 CFA der Strategie Elaborieren	236
5.3.1.2.2 CFA der Strategie Kritisches Prüfen.....	236
5.3.1.2.3 CFA der Strategie Organisieren	237
5.3.1.2.4 CFA der Strategie Wiederholen	237
5.3.1.2.5 Analyse der metakognitiven Strategien.....	238
5.3.1.2.6 CFA der Strategie Anstrengung	239
5.3.1.2.7 CFA der Strategie Aufmerksamkeit.....	240
5.3.1.2.8 CFA der Strategie Zeitmanagement.....	240
5.3.1.2.9 CFA der Strategie Lernumgebung	240
5.3.1.2.10 CFA der Strategie Lernen mit Studienkollegen	241
5.3.1.2.11 CFA der Strategie Literatur	242
5.3.1.2.12 Kurzes Fazit der konfirmatorischen Faktorenanalyse des LIST	242
5.3.2 Zusammenhänge epistemologischer Überzeugungen mit Lernstrategien.....	242
5.3.2.1 Elaborieren	243
5.3.2.2 Kritisches Prüfen	243
5.3.2.3 Organisieren	244
5.3.2.4 Wiederholen	244
5.3.2.5 Planen	244
5.3.2.6 Überwachen.....	245
5.3.2.7 Regulieren	245
5.3.2.8 Anstrengung	246
5.3.2.9 Aufmerksamkeit	246
5.3.2.10 Zeitmanagement	247
5.3.2.11 Lernumgebung	247
5.3.2.12 Literatur	248
5.3.2.13 Zusammenfassung der Regressionsmodelle.....	248
5.4 Diskussion der Regressionsanalysen.....	251
6. Allgemeine Diskussion	260

6.1 Abgrenzung epistemologischer Überzeugung	261
6.2 Erfassung epistemologischer Überzeugungen	270
6.2.1 Itemgestaltung	270
6.2.2 Validitätsaspekte	273
6.2.3 Reaktivität der Itembeantwortung	280
6.3 Normative Sichtweise epistemologischer Überzeugungen.....	284
6.3.1 Produktivität, Korrektheit und Kontextabhängigkeit	284
6.3.2 Normativität und die Entwicklung epistemologischer Überzeugungen.....	289
6.3.3 Methodische Folgerungen der Kontextabhängigkeit	293
6.3.4 Soziale Aspekte epistemologischer Überzeugungen.....	297
6.4 Epistemologische Überzeugungen als Schemata	302
6.5 Eine Latent-State-Trait-Theorie epistemologischer Überzeugungen.....	308
6.5.1 Inhaltliche Darstellung der LST-Theorie epistemologischer Überzeugungen.....	308
6.5.2 Methodische Darstellung der LST-Theorie epistemologischer Überzeugungen ..	312
7. Literaturverzeichnis	321
8. Anhang	346
8.1 Anhang A: Epistemological Belief Questionnaire von Schommer (1990).....	346
8.2 Anhang B: EPI-Fragebogen	349
8.3 Anhang C: Fragebogen zur Erfassung epistemologischer Überzeugungen (FEE)	351
8.4 Anhang D: Inventar zur Erfassung von Lernstrategien im Studium (LIST).....	354

Tabellenverzeichnis

Tabelle 1: Übersicht über die wichtigsten, auf epistemologische Überzeugungen bezogenen Forschungsrichtungen	7
Tabelle 2: Subskalen des EBQ von Schommer (1990) und Anzahl der Items	11
Tabelle 3: Zuordnung der Subskalen zu den Faktoren.....	12
Tabelle 4: Überblick über die Faktorenlösungen (Clarebout et al., 2001).....	13
Tabelle 5: Dimensionen epistemologischer Überzeugungen (Hofer und Pintrich, 1997, Gerber, 2004)	22
Tabelle 6: Übersicht über die Skalen von Gerber (2004).....	24
Tabelle 7: Interkorrelation der Skalen aus Gerber (2004).....	24
Tabelle 8: Skalen zur Messung epistemologischer Überzeugungen nach Elder (2002)	25
Tabelle 9: Skalen zur Messung epistemologischer Überzeugungen (Conley et al., 2004).....	26
Tabelle 10: Interkorrelation der Skalen aus Conley et al. (2004)	27
Tabelle 11: Korrelation der Skalen aus Urhahne und Hopf (2004)	27
Tabelle 12: Interne Konsistenzen und Anzahl der Items der Skalen aus Hofer (2000).....	28
Tabelle 13: A priori Konstrukte nach Moschner und Gruber (2005).....	32
Tabelle 14: Ergebnisse aus Moschner et al. (2005)	35
Tabelle 15: Ergebnisse aus Moschner und Gruber (2005a)	36
Tabelle 16: Ergebnisse aus Moschner und Gruber (2005a)	37
Tabelle 17: Skalen des FEE	37
Tabelle 18: Beispielhafte Aufzählung epistemologischer Ressourcen nach Elby und Hammer (2002)	42
Tabelle 19: Klassifikation akademischer Domänen nach Biglan (1973b).....	48
Tabelle 20: Vergleich verschiedener akademischer Domänen nach Royce (1978).....	49
Tabelle 21: Epistemologische Überzeugungen und metakognitive Komponenten nach Hofer (2004b)	57
Tabelle 22: Ergebnisse von Urhahne und Hopf (2004).....	62
Tabelle 23: Ergebnisse von Phan (2008).....	68
Tabelle 24: Interkorrelation von epistemologischen Überzeugungen und Lernstrategien (Dahl et al., 2005).....	69
Tabelle 25: Multiple Regressionen epistemologischer Überzeugungen auf Lernstrategien (Dahl et al., 2005).....	70
Tabelle 26: Übersicht über die Lernstrategien nach Wild et al. (2009)	73
Tabelle 27: Eignung der Daten zur Berechnung einer Faktorenanalyse mittels des KMO-Koeffizienten	89
Tabelle 28: Beurteilung der Stichprobengröße für eine Faktorenanalyse.....	90
Tabelle 29: Übersicht über die Beurteilungskriterien der verwendeten Fitindizes.....	107
Tabelle 30: Deskriptive Statistik der Altersverteilung in Jahren	135
Tabelle 31: Aufteilung nach Studium bzw. Beruf und Geschlecht.....	136
Tabelle 32: Aufteilung nach Fächern und Geschlecht	136
Tabelle 33: Verteilungsform der EBQ-Items	137
Tabelle 34: Zuordnung der Items zu den Subskalen	139
Tabelle 35: Modellpassung der Subskalen von Schommer	140
Tabelle 36: Korrelationsmatrix der Skala S1	140
Tabelle 37: Korrelationsmatrix der Skala S5	141
Tabelle 38: Parameterschätzungen S6.....	142
Tabelle 39: Parameterschätzungen S10.....	142
Tabelle 40: Faktorladungen der Subskalen	144
Tabelle 41: Vergleich der Faktorenlösung mit den Ergebnissen von Schommer (nach Clarebout et al., 2001).....	145

Tabelle 42: MSA-Werte der Schommer-Items	147
Tabelle 43: Ergebnisse der Parallelanalyse	150
Tabelle 44: Mustermatrix der Promax-Rotation	152
Tabelle 45: Korrelationsmatrix der fünf Faktoren	153
Tabelle 46: Faktorladungen des E/CFA-Modells	154
Tabelle 47: Korrelation der Faktoren in der E/CFA	157
Tabelle 48: Vergleich der Faktorkorrelation von EFA und E/CFA	157
Tabelle 49: Faktorladungen des Faktors (Autoritäres-einfaches Wissen)	159
Tabelle 50: Reliabilitäten des Faktors (Autoritäres Wissen)	159
Tabelle 51: Faktorladungen des Faktors (Existenz von Wahrheit)	160
Tabelle 52: Reliabilitäten der Skala (Existenz von Wahrheit)	161
Tabelle 53: Faktorladungen des Faktors (Integratives Wissen)	162
Tabelle 54: Reliabilitäten der Skala (Integratives Wissen)	162
Tabelle 55: Faktorladungen des Faktor (Lernmethodik)	163
Tabelle 56: Reliabilitäten der Skala (Lernmethodik)	164
Tabelle 57: Faktorladungen des Faktors (Unsicherheit des Wissens)	165
Tabelle 58: Reliabilitäten für die Skala (Unsicherheit des Wissens)	165
Tabelle 59: Mittlere Item-Interkorrelationen und Präzision von Cronbachs α	173
Tabelle 60: Deskriptive Statistik der Altersverteilung in Jahren - Analytestichprobe	183
Tabelle 61: Verteilung der Probanden auf Fächergruppen - Analytestichprobe	183
Tabelle 62: Verteilungsform der EPI-Items – Analytestichprobe	184
Tabelle 63: Deskriptive Statistik der Altersverteilung in Jahren – Validierungsstichprobe ..	185
Tabelle 64: Verteilung der Probanden auf Fächergruppen – Validierungsstichprobe	186
Tabelle 65: Verteilungsform der EPI-Items – Validierungsstichprobe	186
Tabelle 66: Parameterwerte für Modell RN 1	190
Tabelle 67: Reliabilitätsschätzungen der Skala (Reflexive Natur des Wissens)	190
Tabelle 68: Ergebnisse der Skala (Reflexive Natur des Wissens)	191
Tabelle 69: Mustermatrix der zweifaktoriellen Lösung	193
Tabelle 70: Parameterwerte für das Modell S/OWF final	195
Tabelle 71: Reliabilitätsschätzungen der Skalen für Modell S/OWF final	197
Tabelle 72: Ergebnisse des zweifaktoriellen CFA-Modells	198
Tabelle 73: Parameter der Skala (Soziale Komponente des Wissens)	200
Tabelle 74: Reliabilitätsschätzungen der Skala (Soziale Komponente des Wissens)	200
Tabelle 75: Ergebnisse für die Skala (Soziale Komponente des Wissens)	201
Tabelle 76: Parameterwerte des Faktors (Value of Knowledge)	202
Tabelle 77: Reliabilitätsschätzungen für die Skala (Value of Knowledge)	202
Tabelle 78: Fitindizes des Faktors (Value of Knowledge)	203
Tabelle 79: Mustermatrix der zweifaktoriellen Lösung	204
Tabelle 80: Parameter des Faktors (Kulturspezifische Wissenszugänge)	207
Tabelle 81: Reliabilitätsschätzungen für die Skala (Kulturspezifische Wissenszugänge)	207
Tabelle 82: Fitindizes des Faktors (Kulturspezifische Wissenszugänge)	208
Tabelle 83: Parameterwerte für den Faktor (Lernen lernen)	209
Tabelle 84: Reliabilitätsschätzungen für die Skala (Lernen lernen)	209
Tabelle 85: Fitindizes des Faktors (Lernen lernen)	210
Tabelle 86: Modifikationsindizes für das revidierte Gesamtmodell der EPI-Faktoren	213
Tabelle 87: Faktorladung des Gesamtmodells	214
Tabelle 88: Geschätzte Korrelation der Faktoren im Gesamtmodell	215
Tabelle 89: Mustermatrix der ML-EFA	215
Tabelle 90: Korrelationsmatrix der Faktoren	216
Tabelle 91: Mustermatrix der PAF-EFA	217
Tabelle 92: Korrelationsmatrix der Faktoren	217

Tabelle 93: Vergleich der EPI- und FEE-Faktoren.....	221
Tabelle 94: Schritte bei der Prüfung faktorieller Invarianz.....	224
Tabelle 95: Übersicht über die Reliabilitäten der EPI-Skalen	229
Tabelle 96: Deskriptive Statistik der Altersverteilung in Jahren	232
Tabelle 97: Aufteilung der Probanden auf die Fächergruppen	233
Tabelle 98: Verteilungsform der LIST Items	233
Tabelle 99: EFA der metakognitiven Strategien	238
Tabelle 100: Korrelation der drei Faktoren.....	238
Tabelle 101: Modifikationsindizes für die Strategie Lernen mit Studienkollegen	241
Tabelle 102: Regression EPI auf Elaborieren	243
Tabelle 103: Regression EPI auf Kritisches Prüfen	243
Tabelle 104: Regression EPI auf Organisation	244
Tabelle 105: Regression Wiederholen auf EPI	244
Tabelle 106: Regression Planen auf EPI	245
Tabelle 107: Regression Überwachen auf EPI.....	245
Tabelle 108: Regression Regulieren auf EPI	246
Tabelle 109: Regression Anstrengung auf EPI	246
Tabelle 110: Regression Aufmerksamkeit auf EPI	247
Tabelle 111: Regression Zeitmanagement auf EPI	247
Tabelle 112: Regression Lernumgebung auf EPI	248
Tabelle 113: Regression Literatur auf EPI.....	248
Tabelle 114: Zusammenfassung der Regressionsergebnisse für die kognitiven und metakognitiven Lernstrategien.....	249
Tabelle 115: Zusammenfassung der Regressionsergebnisse für die ressourcenbezogenen Lernstrategien.....	249
Tabelle 116: Klassifikation der Effektstärken nach Cohen (1988) für kognitive und metakognitive Lernstrategien.....	249
Tabelle 117: Klassifikation der Effektstärken nach Cohen (1988) für die ressourcenbezogenen Lernstrategien.....	250
Tabelle 118: Dimensionen nach Hofer und Pintrich (1997) Erfüllung des Rechtfertigungskriteriums	264
Tabelle 119: Dimensionen des revidierten EPI-Fragebogens sowie die Erfüllung des Rechtfertigungskriteriums	267
Tabelle 120: Tabelle mit den tendenziellen Änderungen aus Kienhues et al. (2008).....	290
Tabelle 121: Prozentualer Anteil der Probanden mit spontanen Äußerungen über epistemologische Überzeugungen (Ferguson et al., 2012)	305
Tabelle 122: Nach Dimensionen aufgeteilte Prozentsätze aller abgegebenen Äußerungen (Ferguson et al., 2012).....	306

Abbildungsverzeichnis

Abbildung 1: Dimensionale Struktur von Schommers Modell epistemologischer Überzeugungen..... 8

Abbildung 2: Embedded-Belief-System von Schommer-Aikins (2004) 16

Abbildung 3: Angenommene Interaktion im EBS 18

Abbildung 4: Rolle epistemologischer Überzeugungen in der Gestaltung von Lehr-Lern-Situationen (Gruber & Moschner, 2003)..... 31

Abbildung 5: TIDE-Framework von Muis et al. (2006), dargestellt für die drei Domänen A, B und C 50

Abbildung 6: Wirkmodell von Hofer (2001) 53

Abbildung 7: Kognitive Pyramide nach Kitchener (1983) 55

Abbildung 8: Das EBS von Schommer-Aikins 61

Abbildung 9: Pfadmodell von Cano (2005) 63

Abbildung 10: Resultierendes Pfadmodell mit signifikanten Pfaden nach Cano (2005)..... 65

Abbildung 11: Pfadmodell von Phan (2006)..... 66

Abbildung 12: Darstellung eines CFA-Modells mit einer latenten Variablen..... 96

Abbildung 13: Darstellung eines CFA-Modells mit zwei latenten Variablen 96

Abbildung 14: CFA-Modell mit zwei latenten Variablen und Kreuzladungen zu allen manifesten Variablen..... 98

Abbildung 15: CFA-Modell mit korrelierten Messfehlern 99

Abbildung 16: CFA-Modell zur Berechnung der Reliabilität..... 121

Abbildung 17: Nomologisches Netz (in Anlehnung an Hartig et al., 2008)..... 125

Abbildung 18: Schematische Darstellung des Messmodells und Strukturmodells..... 126

Abbildung 19: Strukturgleichungsmodell zur Konstruktvalidierung..... 129

Abbildung 20: Strukturgleichungsmodell mit drei latenten Variablen und Kovarianzen..... 132

Abbildung 21: Scree-Plot der Schommer Subskalen 143

Abbildung 22: Scree-Plot der Schommer Items..... 151

Abbildung 23: Scree-Plot der Skala Absolutes Wissen 192

Abbildung 24: Scree-Plot der Faktorenanalyse der Skala Geschlechtsspezifische Wissenszugänge 204

Abbildung 25: CFA-Modell mit zwei Methodenfaktoren für die Skala Geschlechtsspezifische Wissenszugänge (unterstrichene Items haben ein negatives Regressionsgewicht) 205

Abbildung 26: Ergebnisse aus Maggioni et al. (2004)..... 291

Abbildung 27: Messwert als latenter Trait-Wert und latentes State-Residuum..... 313

Abbildung 28: Single-Trait-Multi-State Modell 317

1. Überblick zur Forschung

1.1 Einführung

Die vorliegende Arbeit befasst sich mit epistemologischen Überzeugungen. Epistemologische Überzeugungen sind ein Konstrukt in der Psychologie, welches die Annahmen eines Individuums über die Struktur des Wissens sowie den Prozess des Wissenserwerbs bezeichnet (Hofer, 2002). Die Forschung über epistemologische Überzeugungen erlebte in der letzten Dekade einen Aufschwung, was zahlreiche Überblicksartikel (Hofer & Pintrich, 1997) und Einzelbeiträge (Muis, Bendixen & Haerle, 2006, Mason & Bromme, 2010), Sonderbände von Zeitschriften (Hofer, 2004a, Limón, 2006a) sowie Sammelbände (Hofer & Pintrich, 2002, Khine, 2008, Bendixen & Feucht, 2010) belegen. Mit dem Aufschwung der Forschung ging auch eine erhebliche Weiterentwicklung der theoretischen Konzeption epistemologischer Überzeugungen einher.

Hauptsächlich im Vordergrund steht dabei ein bestimmter Zugang zum Thema epistemologische Überzeugungen, nämlich den persönlichkeitspsychologischen Ansatz, wie er in der Pädagogischen Psychologie vorherrschend ist. Die Forschungsfragen, die behandelt werden, zielen auf die Validierung zweier Instrumente zur Erfassung epistemologischer Überzeugungen ab. Diese Frage nach der Validierung zweier Messinstrumente führte schließlich zu einem Entwurf für die theoretische und methodologische Neukonzeption epistemologischer Überzeugungen.

Das erste Kapitel beginnt mit einem Überblick über die Geschichte der Erforschung epistemologischer Überzeugungen. Anschließend folgt eine umfassende Übersicht über die persönlichkeitspsychologischen Ansätze epistemologischer Überzeugungen. Die Darstellung verschiedener Herangehensweisen beschränkt sich jedoch nicht nur auf eine Zusammenfassung der persönlichkeitspsychologischen Modelle. Es werden auch andere Herangehensweisen, wie sie beispielsweise in der Fachdidaktik einzelner Fächer zu finden sind, vorgestellt. Außerdem werden Modelle aufgeführt, die epistemologische Überzeugungen im instruktionalen Umfeld bzw. dem kognitiven System des Menschen verorten. Diese alternativen Herangehensweisen werden aufgeführt, da sich die vorgestellten Theorien und besonderen Methoden dieser Ansätze im Verlauf der Argumentation als wichtig erweisen werden.

Neben der Vorstellung von Theorien und Modellen epistemologischer Überzeugungen behandelt das erste Kapitel aber auch die pädagogische Relevanz epistemologischer Überzeugungen. Hierzu werden empirische Belege angeführt, welche die Wirkung epistemologischer Überzeugungen auf Lernstrategien aufzeigen. Der Abschluss der Einführung stellt die methodologische Sichtweise vor, welche der Konzeption epistemologischer Überzeugungen in den hier durchgeführten empirischen Arbeiten zugrunde liegt. Im Anschluss folgen im zweiten Kapitel die Formulierung der Forschungsfragen sowie im dritten Kapitel eine ausführliche Darstellung der Methoden, die zur Beantwortung der Forschungsfragen eingesetzt werden.

Die durchgeführten Untersuchungen gliedern sich in zwei Teile. In der ersten Untersuchung im vierten Kapitel werden die Eigenschaften des wahrscheinlich am häufigsten eingesetzten Fragebogens zur Erfassung epistemologischer Überzeugungen untersucht. In der zweiten Untersuchung im fünften Kapitel wird ein neuerer und bisher nicht oft angewendeter Fragebogen aus dem deutschsprachigen Raum untersucht. Mithilfe dieses Fragebogens werden die Beziehungen der mit ihm erfassten Dimensionen epistemologischer Überzeugungen zu Lernstrategien analysiert. Den jeweiligen Untersuchungen schließen sich Diskussionen über die Ergebnisse an.

Den Abschluss der Arbeit im sechsten Kapitel bildet eine zusammenfassende Diskussion der Befunde der durchgeführten empirischen Untersuchungen. Dabei werden die Befunde sowohl aus inhaltlicher als auch methodischer Sichtweise diskutiert. Die abschließende Diskussion mündet in einer neuen psychologischen und methodischen Theorie epistemologischer Überzeugungen, die einen Ansatzpunkt für weitergehende Forschung darstellen kann.

1.2 Definition des Begriffs „epistemologische Überzeugungen“ und geschichtlicher Abriss des Konzepts

Zu Beginn wird eine Definition epistemologischer Überzeugungen aufgeführt, die für die weitere Arbeit zugrunde gelegt werden kann, daran anschließend erfolgen ein kurzer historischer Abriss der Forschungen sowie eine Charakterisierung der wichtigsten Forschungsströmungen.

Unter epistemologischen Überzeugungen („*epistemological beliefs*“) werden im Folgenden die Annahmen einer Person über die Natur des Wissens und den Prozess des Wissenserwerbs verstanden. Epistemologische Überzeugungen bezeichnen also subjektive Vorstellungen über

die Objektivität, die Richtigkeit, die Aussagekraft oder die Herkunft von Wissen (vgl. Gruber & Stamouli, 2009).

Bei epistemologischen Überzeugungen handelt es sich nicht um ein einheitlich definiertes Konstrukt, worauf schon seine vielfältigen Bezeichnungen hindeuten. Beispielsweise finden sich die Bezeichnungen *epistemological beliefs* (Schommer 1990), *personal epistemology* (Hofer und Pintrich, 1997), *epistemological resource* (Hammer & Elby, 2002), *epistemological position* (Perry, 1970), *epistemological standard* (Ryan, 1984), *epistemological reflection* (Baxter Magolda, 1992) oder aber *ways of knowing* (Belenky, Clinchy, Goldberger & Tarule, 1986) und *reflective judgment* (King, Kitchener, Davison, Parker & Wood, 1983). Dabei besitzt jeder der benannten Ansätze sowohl in konzeptueller als auch in methodischer Hinsicht einen eigenen Schwerpunkt. Gemeinsames Merkmal der genannten Ansätze sind die kognitiven Prozesse eines Individuums, die bei der Beurteilung von Wissen hinsichtlich seiner Struktur und seiner Gültigkeit eine Rolle spielen.

Grundsätzlich lassen sich vier Forschungsrichtungen unterscheiden, nämlich eine entwicklungspsychologische Richtung, eine weitere persönlichkeitspsychologische Richtung, die epistemologische Überzeugungen als ein System unabhängiger Eigenschaften ansieht, der fachdidaktische Ansatz sowie weitere Ansätze, die sich jeweils mit einem besonderen Aspekt epistemologischer Überzeugungen beschäftigen. Der fachdidaktische Ansatz ist als Oberbegriff zu verstehen, denn er gliedert sich wiederum in verschiedene Ansätze auf, die sich auf die Didaktik eines bestimmten Fachs beziehen. Darüber hinaus gibt es in neuerer Zeit weitere Sichtweisen auf epistemologische Überzeugungen, die sich keinem der Ansätze zuordnen lassen (Hofer, 2001, Priemer, 2006). Hierzu zählen z. B. das TIDE-Modell von Muis et al. (2006), das sich auf die Entwicklung und Domänenspezifität epistemologischer Überzeugungen bezieht, oder das Modell von Buehl und Alexander (2006), das sich ebenfalls mit der Domänenspezifität epistemologischer Überzeugungen auseinandersetzt, sowie die psychoepistemologischen Profile von Royce (1978), welche die kognitive Struktur und Domänenspezifität beschreiben.

Ein Beispiel aus dem entwicklungspsychologischen Ansatz ist das *Epistemological-position-Modell* von Perry (1970). Der konzeptionelle Schwerpunkt liegt hierbei im Bereich der individuellen Entwicklung epistemologischen Denkens und der methodische Schwerpunkt der Erkenntnisgewinnung liegt in der Anwendung von Interviews. Die gewonnenen Erkenntnisse

werden schließlich zu einem ontogenetischen Stufenmodell zusammengefasst, das die Entwicklung des epistemologischen Denkens im Verlauf der akademischen Ausbildung beschreibt. Das *Reflective-judgment*-Modell von King et al. (1983) knüpft in wesentlichen Teilen an das Perrys Modell an, indem es dessen Stufenmodell ausbaut, aber zum Teil einen anderen methodischen Zugang wählt. Die Erfassung der epistemologischen Prozesse geschieht allerdings nicht mehr durch Interviews, sondern durch Auswerten der Antworten, welche die Versuchspersonen bei der Bearbeitung schlecht strukturierter Probleme aus verschiedenen Wissensbereichen geben. Die grob umrissenen Modelle sind bekannte Beispiele für eine *entwicklungspsychologische* Herangehensweise an das Thema epistemologische Überzeugungen. Der Unterschied zwischen den beiden Modellen besteht im methodischen Zugang der Erfassung epistemologischer Überzeugungen.

Kennzeichnend für die entwicklungspsychologische Betrachtungsweise ist die Verwendung ontologischer Stufenmodelle. Dieses Vorgehen knüpft an das Werk von Piaget an, der die Entwicklung eines Individuums vom Säuglingsalter bis zum Erwachsenenalter durch verschiedene, jeweils qualitativ unterschiedlich geartete Entwicklungsstufen beschreibt.¹ Epistemologische Überzeugungen werden hier als ein eindimensionales Konstrukt aufgefasst. Die auf jeder Stufe beschriebenen kognitiven Prozesse verändern sich demnach nicht in ihrer Ausprägung, sondern sind in jeder Stufe gänzlich anderer Natur. So beschreibt z. B. die erste Stufe von Perrys Modell, *Dualism*, eine Art des Denkens, die von den Kategorien *richtig* oder *falsch* ausgeht und Autoritäten die Kompetenz zuschreibt, die jeweils richtige Aussage zu kennen und diese auch an andere vermitteln zu können. Die zweite Entwicklungsstufe, *Multiplicity*, ist gekennzeichnet durch die Erkenntnis, dass die Unterscheidung in *richtig* oder *falsch* zu eng ist und es durchaus Unsicherheiten in der Beurteilung einer Aussage geben kann, wobei diese aber letztendlich durch die Autoritäten beigelegt werden. Auf der *Multiplicity*-Stufe wird also eine Erweiterung der Sichtweise eingeführt, wie sie in der *Dualism*-Stufe nicht vorhanden war. In den Stufenmodellen beginnen Individuen demzufolge mit ihrer Entwicklung immer auf der ersten Stufe und erreichen fortgeschrittene Stufen nur dann, wenn die vorausgehenden Stufen durchlaufen wurden. Das heißt, eine Stufe zu überspringen, ist nicht möglich, die Entwicklung muss nicht immer zwangsweise bis zur letzten Stufe voranschreiten, sondern sie kann auf jeder Stufe enden. Die Entwicklungsstufen führen zu einem nicht mehr zu steigernden Endzustand. In der neueren Entwicklungspsychologie sind diese Stufenmodelle allerdings wegen der starren Entwicklungssequenz verworfen worden (vgl. Montada,

¹ Jean Piaget wird heute überwiegend als Entwicklungspsychologe angesehen. Tatsächlich resultierte seine Beschäftigung mit der kindlichen Entwicklung aber aus primär philosophischem Interesse (Chapman, 1988).

2008).² So lässt sich an dieser Stelle zunächst einmal festhalten, dass sich die Klasse der entwicklungspsychologischen Theorien über epistemologische Überzeugungen durch die Verwendung von Stufenmodellen kennzeichnen lässt.

Die Forschung im Rahmen der entwicklungspsychologischen Sichtweise lässt sich noch weiter differenzieren. Wie zuvor bei der groben Skizzierung des *Epistemological-position-* und *Reflective-judgement-*Modells angemerkt, unterscheiden sich diese in der methodischen Herangehensweise. Während Perry durch die Verwendung von Interviews und deren qualitativer Auswertung zu seinem Modell gelangte, wählten King et al. (1983) als Zugang zur Erfassung epistemologischer Überzeugungen die Bearbeitung schlecht strukturierter Probleme aus verschiedenen Wissensbereichen durch die Probanden. Die Auswertung des Bearbeitungsprozesses erfolgt hier allerdings ebenfalls mithilfe qualitativer Methoden. Die Verwendung qualitativer Auswertungsmethoden ist ein weiteres gemeinsames Merkmal der entwicklungspsychologischen Modelle.

Einen gänzlich anderen Zugang stellt die Forschung dar, in der epistemologische Überzeugungen als ein System von Eigenschaften angesehen werden. In diesem wird weniger die Ontogenese epistemologischer Überzeugungen untersucht als vielmehr deren interne Struktur und zum Teil ihr Bezug zum Lernen. Zu den ersten Ansätzen dieser Forschungsrichtung gehört das *Epistemological-belief-*Modell von Schommer (1990). Während die beiden zuvor genannten Stufenmodelle die Eindimensionalität epistemologischen Denkens betonen und den Übergang zwischen den Entwicklungsstufen als qualitativen Sprung ansehen, konzipierte Schommer in ihrem Modell epistemologische Überzeugungen als ein System unterschiedlicher Dimensionen, welche sich in ihrer Ausprägung gleichzeitig und quantitativ statt qualitativ verändern können. Bei diesem Modell kommt hinzu, dass die Beziehung epistemologischer Überzeugungen zu Lernvorgängen dadurch verankert wird, dass Überzeugungen im Hinblick auf das Lernverhalten als Teil epistemologischer Überzeugungen angesehen werden. In methodischer Hinsicht typisch ist die Verwendung mehrdimensionaler Modelle, ähnlich den mehrdimensionalen Persönlichkeitsmodellen aus der Differenziellen Psychologie (vgl. Asendorpf, 1999). Das Modell von Schommer (1990) besteht aus den (a priori ange-

² Dabei ist die Zuordnung von Perrys Modell zu den Stufenmodellen, insbesondere im Hinblick auf die Kritik der Stufenmodelle (Montada, 2008), fraglich. Perrys *epistemological positions* sind nicht notwendig einer fest vorgegebenen Entwicklungssequenz unterworfen, vielmehr können einzelne Positionen übersprungen werden und auch eine Rückkehr auf vorhergehende Stufen ist möglich (Schommer, 1994). Allerdings sind die nachfolgenden, auf Perry aufbauenden Modelle als Stufenmodell konzipiert und später wurde Perrys Modell als Stufenmodell kategorisiert (vgl. Hofer & Pintrich, 1997).

nommenen) Dimensionen *Certainty of knowledge*, *Source of knowledge*, *Structure of knowledge*, *Control of Learning* und *Quick learning*. Die letzten beiden Dimensionen beziehen sich auf Lernprozesse. Auf Schommer geht die Verwendung von Fragebögen zur Erfassung epistemologischer Überzeugungen zurück. Die Verwendung von Fragebogenverfahren vereinfacht die Datenaufnahme erheblich, erlaubt die ökonomische Untersuchung größerer Gruppen und gestattet es so, Fragen zu untersuchen, die nicht durch aufwendige Interviewstudien zu beantworten sind. Allerdings geht dabei die Entwicklungssequenz verloren (Krettenauer, 2005). Schommers Modell ist der Prototyp für eine Reihe weiterer Modelle, wobei jedoch in den meisten Fällen der von Schommer entwickelte Fragebogen als Ausgangspunkt verwendet wurde. Ein weiteres Modell dieser Klasse von Ansätzen stellt das Modell epistemologischer Überzeugungen als Persönliche Theorien von Hofer und Pintrich (1997) dar, das von den vier Dimensionen *Sicherheit des Wissens*, *Komplexität des Wissens*, *Wissensquelle* und *Wissensbegründung* ausgeht. Die Modelle von Schommer (1990) sowie Hofer und Pintrich (1997) sind die prominentesten Vertreter der Sichtweise, die epistemologische Überzeugungen als ein System unabhängiger Eigenschaften betrachtet. Da diese Sichtweise in der angewandten Methodik stark von der Differenziellen bzw. Persönlichkeitspsychologie geprägt ist, findet sich hierfür auch die Bezeichnung „persönlichkeitspsychologische Sichtweise epistemologischer Überzeugungen“. Neben dieser Bezeichnung wird gelegentlich auch vom „pädagogisch-psychologischen Ansatz“ gesprochen, da oftmals die Beziehung zum Lernen mithilfe von Fragebögen erfasst und der entsprechenden Methodik ausgewertet wird.

Einen weiteren, sich mit epistemologischen Überzeugungen befassenden Forschungszweig stellen die Ansätze dar, die in der Didaktik einzelner Fächer zu finden sind. Diese beschäftigen sich nicht generell mit epistemologischen Überzeugungen, sondern mit dem Zusammenspiel von epistemologischen Überzeugungen und fachdidaktischen Überlegungen. Am bekanntesten sind die Ansätze aus der Didaktik der Physik, wie z. B. die epistemologischen Ressourcen von Hammer und Elby (2002), sowie aus der Mathematikdidaktik (Schoenfeld, 1985). Zu den bereits genannten Ansätzen existieren weitere, die sich nicht in das bis hierhin dargestellte Schema einordnen lassen, dies sind die o. g. alternativen Ansätze (Hofer, 2001). Darunter fällt z. B. das Modell von Hofer (2004b), in dem epistemologische Überzeugungen als metakognitive Prozesse aufgefasst werden, das TIDE-Modell von Muis et al. (2006), das Modell von Buehl und Alexander (2006) sowie der schon ältere Ansatz der psychoepistemologischen Profile von Royce (1978). Diese Modelle greifen einzelne Aspekte auf, die bisher

nicht oder nur unzureichend berücksichtigt wurden. Tabelle 1 gibt eine Übersicht über die vorgestellten theoretischen Modelle epistemologischer Überzeugungen.

Tabelle 1: Übersicht über die wichtigsten, auf epistemologische Überzeugungen bezogenen Forschungsrichtungen

Forschungsrichtung	Beispielhafte prominente Vertreter
Entwicklungspsychologie	<i>Epistemological positions</i> von Perry (1968), <i>Reflective-Judgement</i> -Modell von King et al. (1983)
Pädagogische Psychologie bzw. persönlichkeitspsychologische Modelle	Modell von Schommer (1990), epistemologische Überzeugungen als Persönliche Theorien nach Hofer und Pintrich (1997)
Fachdidaktik	Epistemologische Ressourcen von Hammer und Elby (2002)
Alternative Ansätze	TIDE-Modell von Muis et al. (2006), Modell von Buehl und Alexander (2006), psychoepistemologische Profile von Royce (1978)

1.3 Persönlichkeitspsychologische Modelle epistemologischer Überzeugungen

Im Folgenden werden diejenigen Ansätze vorgestellt, auf die in den sich anschließenden Kapiteln Bezug genommen wird. An erster Stelle stehen hier die persönlichkeitspsychologischen Modelle, die in diesem Unterkapitel ausführlich vorgestellt werden. Von besonderer Relevanz sind der Ansatz von Schommer (1990) und die darauf aufbauenden Arbeiten, die theoretischen Erweiterungen von Schommer-Aikins (2004) sowie das Modell von Gruber und Moschner (2003). Diese Arbeiten sind grundlegend für die im Rahmen der vorliegenden Arbeit durchgeführten empirischen Untersuchungen. Ebenso ausführlich vorgestellt wird das Modell von Hofer und Pintrich (1997), vor allem wegen seiner Bedeutung als weiterer Vertreter der persönlichkeitspsychologischen Modelle.

1.3.1 Schommers *Epistemological-Belief-Modell*

Marlene Schommer (1990) konzipierte epistemologische Überzeugungen in ihrem Ansatz als die Überzeugungen über die Struktur des Wissens und den Prozess des Wissenserwerbs, die ein System „mehr oder weniger“³ unabhängiger Dimensionen bilden. Dieses System besteht aus den folgenden fünf Dimensionen, beschrieben werden jeweils die beiden Pole:

³ Obwohl dem äußerst unscharfen Ausdruck „mehr oder weniger“ sowohl aus der Sicht einer klaren Definition als auch aus der Sicht der verwendeten Methodik Kritik entgegengebracht werden kann, soll er an dieser Stelle benutzt werden, weil Schommer diesen Sprachgebrauch in ihren Publikationen durchgängig benutzt und das Hauptziel dieses Kapitels die Beschreibung ihres Modells ist.

- *Source of knowledge*: „Wissen wird durch eine allwissende Autorität vermittelt“ bis „Wissen wird durch objektive und subjektive Mittel erschlossen“.
- *Certainty of knowledge*: „Wissen ist absolut“ bis „Wissen entwickelt sich ständig weiter“.
- *Structure of knowledge/Organization of knowledge*: „Wissen ist in kleine Bestandteile unterteilt“ bis „Wissensbestandteile sind eng miteinander verbunden und integriert“.
- *Innate ability/Control of learning*: „Fähigkeit zum Lernen ist genetisch vorherbestimmt“ bis „Fähigkeit zum Lernen ist durch Übung angeeignet“
- *Quick learning/Speed of Learning*: „Lernen geschieht schnell oder überhaupt nicht“ bis „Lernen ist ein gradueller Prozess“.

Diese fünf Dimensionen lassen sich jeweils zu den beiden übergeordneten Dimensionen *Beliefs about knowledge* und *Beliefs about knowledge acquisition* zusammenfassen (vgl. Abbildung 1). Dabei bilden die Dimensionen *Source of knowledge*, *Certainty of knowledge* und *Structure of knowledge* die übergeordnete Dimension *Beliefs about knowledge*, die Dimensionen *Control of learning* und *Speed of Learning* bilden die übergeordnete Dimension *Beliefs about knowledge acquisition*.

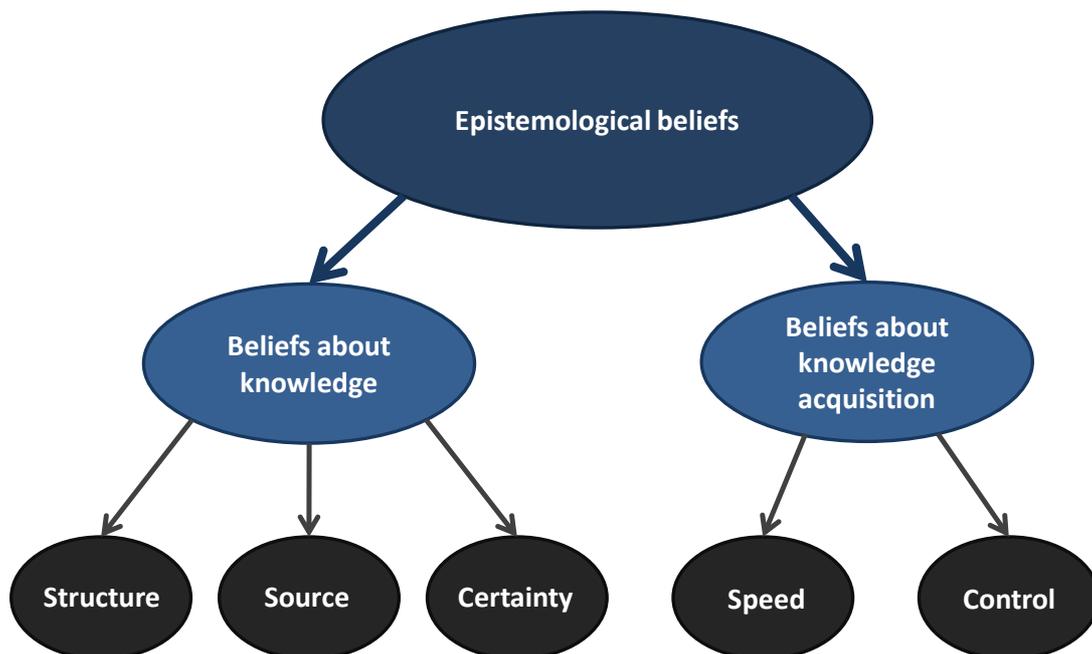


Abbildung 1: Dimensionale Struktur von Schommers Modell epistemologischer Überzeugungen

Schommer gibt mit ihrem System nicht nur die dimensionale Struktur epistemologischer Überzeugungen an. In einer Erweiterung des ursprünglichen Ansatzes (Schommer-Aikins, 2004) werden zusätzlich Merkmale epistemologischer Überzeugungen aufgeführt. Die wesentlichen Merkmale des *Epistemological-belief*-Modells lassen sich durch folgende Punkte beschreiben:

- Überzeugungen zum Lernen sind ein fester Bestandteil des Systems epistemologischer Überzeugungen.
- Die Überzeugungen zum Wissen, zum Prozess des Wissenserwerbs sowie zum Lernen bilden zwar ein System, sind aber voneinander getrennt zu betrachten.
- Epistemologische Überzeugungen können sich unabhängig voneinander (asynchron) entwickeln.
- Es gibt ein Bedürfnis nach der Balance epistemologischer Überzeugungen.
- Das *Epistemological-belief*-Modell führt eine einheitliche Nomenklatur epistemologischer Überzeugungen ein.
- Es gibt die Möglichkeit einer quantitativen Erfassung epistemologischer Überzeugungen.

Diese Merkmale des *Epistemological-belief*-Modells und die sich daraus ergebenden Implikationen und Kritikpunkte werden im Folgenden dargestellt. Überzeugungen zum Lernen werden als Teil epistemologischer Überzeugungen konzipiert. Dies ist jedoch in der Sichtweise einer erkenntnistheoretischen Abgrenzung epistemologischer Überzeugungen zu anderen Konstrukten umstritten (vgl. Rozendaal, de Brabander & Minnaert, 2001). Im *Epistemological-belief*-Modell wird lediglich die Verbindung epistemologischer Überzeugungen zum Lernen referiert, was Schommer dazu nutzt, die Integration der Überzeugungen zum Lernen in das System epistemologischer Überzeugungen zu rechtfertigen. Schommer führt dazu die Befunde von Schoenfeld (1985) an, die aus dem Bereich der epistemologischen Überzeugungen zur Mathematik stammen. Danach sind Schüler der Überzeugung, dass nur Autoritäten Mathematik verstehen können, die über eine entsprechende Begabung verfügen. Ebenso gibt es die Überzeugung, dass nur diese Autoritäten die Beweise mathematischer Sachverhalte führen können und dass mathematische Probleme schnell oder gar nicht gelöst werden. Diese Befunde beziehen sich auf die Dimensionen *Source of knowledge*, *Innate ability* und *Quick learning*. Die Fähigkeit zum Lernen steht mit epistemologischen Überzeugungen derart in Verbindung, dass nur begabte Menschen in der Lage sind, die mathematischen Sachverhalte richtig und schnell zu verstehen und weiterzugeben. Nur diese Menschen können entsprechende mathematische Kompetenzen erlernen und zu Autoritäten auf dem Gebiet der Mathematik werden, um anschließend das Wissen an andere zu vermitteln. Weiter bezieht sich Schommer auf Befunde von Dweck und Leggett (1998), denen zufolge Kinder, die glauben weniger begabt zu sein als andere, auch weniger Zeit in das Lösen von Problemen und andere Lernaktivitäten investieren. Ebenso werden Befunde von Ryan (1984) aufgeführt, der gezeigt hat, dass Lernende, die an ein einfach strukturiertes Wissen glauben, Verständnis des Inhalts mit der Repetition von bloßen Fakten verwechseln. Aufgrund dieser Befunde betrachtet

Schommer die Überzeugungen zum Lernen als einen Teil der epistemologischen Überzeugungen.

Weiterhin wird im Epistemological-belief-Modell postuliert, dass die Überzeugungen zur Struktur des Wissens (*Simplicity of knowledge, Certainty of knowledge*), zum Prozess des Wissenserwerbs (*Source of knowledge*) sowie die Überzeugungen zum Lernen (*Controll of learning, Quick learning*) voneinander getrennt sind, jedoch ein System mit „mehr oder weniger“ unabhängigen Dimensionen bilden. Die Konzeption als „mehr oder weniger“ voneinander unabhängiger Dimensionen erlaubt eine asynchrone Entwicklung. Sind die Dimensionen epistemologischer Überzeugungen orthogonal, kann sich jede Dimension unabhängig von der anderen entwickeln. Liegen korrelierte Dimensionen vor, so bewirkt die Entwicklung einer Dimension eine gleichzeitige Entwicklung der mit ihr korrelierten Dimensionen. Damit soll zum Ausdruck gebracht werden, dass sich epistemologische Überzeugungen unabhängig voneinander entwickeln können, dies aber nicht müssen.

Das Bedürfnis nach Balance trägt der Tatsache Rechnung, dass extreme Ausprägungen epistemologischer Überzeugungen problematisch sein können (Schommer-Aikins, 2002). Wird z. B. jedes Wissen als tentativ angesehen, kann dies in den Fällen kritisch werden, in denen bestimmte Sachverhalte als erwiesen betrachtet werden können. Schommers Konzeption der Balance sieht epistemologische Überzeugungen als Häufigkeitsverteilungen (Schommer, 1994). So wird z. B. ein Großteil des Wissens von einer Person als tentativ, der verbleibende Rest des Wissens als sicher betrachtet.

Durch die einheitliche Nomenklatur der Dimensionen wird eine eindeutige Kommunikation über epistemologische Überzeugungen ermöglicht. Diese fünf Dimensionen wurden von Schommer a priori postuliert und mit Namen versehen. Allerdings ist die Zahl von fünf Dimensionen nicht durchgängig in empirischen Arbeiten von Schommer belegt worden. So sprechen Schommer, Crouse und Rhodes (1992) nur von den vier Dimensionen *Source of Knowledge, Certainty of knowledge, Innate ability* und *Quick Learning*. Hinzu kommt, dass in den Publikationen nicht nur die Zahl der angenommenen Dimensionen schwankt, sondern auch deren Beschreibung unterschiedlich ausfällt (zusammenfassend s. Clarebout, Elen, Luyten & Bamps, 2001). Im Rahmen der vorliegenden Arbeit wird die oben aufgeführte Nomenklatur und Beschreibung aus Schommer (1994) verwendet.

Zur Messung und damit einhergehend der quantitativen Erfassung epistemologischer Überzeugungen wurde der *Epistemological Belief Questionnaire* (EBQ) entwickelt. Dieser Fragebogen wurde in einer der in der vorliegenden Studie durchgeführten empirischen Untersuchungen genutzt und ist darüber hinaus einer der am häufigsten eingesetzte Fragebogen zur Erfassung epistemologischer Überzeugungen. Daher wird der EBQ im Folgenden ausführlich vorgestellt. Der EBQ besteht aus 63 Items, die in Anhang A aufgeführt sind. Die Items des Fragebogens sind 12 a priori angenommenen Subskalen⁴ zugeordnet, die Subskalen mit der dazugehörigen Anzahl der Items sind in Tabelle 2 zu finden. Diese Subskalen werden wiederum den fünf Faktoren des Schommer'schen Modells zugeordnet (siehe Tabelle 3). In der praktischen Anwendung werden die einer Subskala zugeordneten Items summiert und anschließend auch die einem Faktor zugeordneten Subskalen summiert, um so die Ausprägung einer epistemologischen Überzeugung zu bestimmen.

Zur Validierung der faktoriellen Struktur wurde der entwickelte Fragebogen in einer Reihe von Untersuchungen eingesetzt, die auf faktorenanalytischer Basis ausgewertet wurden (Schommer, 1990, 1993a, Schommer et al. 1992). Die Faktorenanalysen wurden allerdings nicht auf Basis der einzelnen Items, sondern auf Basis der Subskalen durchgeführt. Dabei konnten die a priori angenommenen fünf Dimensionen nicht bestätigt werden. Es ergaben sich 3- und 4-faktorielle Lösungen. Tabelle 4 bietet eine von Clarebout et al. (2001) erstellte Übersicht über die verschiedenen Faktorenlösungen aus den Arbeiten der Arbeitsgruppe um Schommer.

Tabelle 2: Subskalen des EBQ von Schommer (1990) und Anzahl der Items

Nr.	Skala	Anzahl Items
S1	Success is unrelated to hard work	4
S2	Avoid ambiguity	5
S3	Seek single answers	11
S4	Avoid Integration	8
S5	Depend on authority	4
S6	Ability to learn is innate	4
S7	Don't criticize authority	6
S8	Knowledge is certain	6
S9	Learning is quick	5
S10	Can't learn how to learn	5
S11	Learn first time	3
S12	Concentrated effort is a waste of time	2

⁴ Diese Subskalen werden auch als Subsets bezeichnet. In der vorliegenden Arbeit werden beide Ausdrücke nebeneinander benutzt.

Tabelle 3: Zuordnung der Subskalen zu den Faktoren

Faktor	Subskala	Subskala Nr.
Simple knowledge	Seek single answers	S3, S4
	Avoid integration	
Certain knowledge	Avoid ambiguity	S2, S8
	Knowledge is certain	
Omniscient authority	Depend on authority	S5, S7
	Don't criticize authority	
	Success is unrelated to hard work	
Innate ability	Ability to learn is innate	S1, S6, S10
	Can't learn how to learn	
	Learning is quick	
Quick learning	Learn first time	S9, S11, S12
	Concentrated effort is a waste of time	

In der Übersicht ist die Zuordnung der einzelnen Subskalen auf den jeweiligen Faktoren abgebildet, in zwei Fällen wird neben der Zuordnung auch die numerische Höhe der Faktorladung angegeben. Zur Bestimmung der Zahl bedeutsamer Faktoren wurde das Kaiser-Kriterium angewandt, nach dem die Eigenwerte der bedeutsamen Faktoren > 1 sein sollen. In der Untersuchung von Schommer et al. (1992) erzielte der vierte Faktor nur einen Eigenwert von .95. Da jedoch in einer zweiten Untersuchung eine konfirmatorische Faktorenanalyse eine bessere Anpassung der 4-faktoriellen Lösung ergab und diese auch die Lösung aus der Studie von Schommer (1990) replizierte, wurde im weiteren Verlauf Bezug auf die 4-faktorielle Lösung genommen. In der Studie von Schommer (1993a) ergab sich bei strikter Anwendung des Kaiser-Kriteriums eine 3-faktorielle Lösung. Da aber der vierte Faktor einen Eigenwert von .98 aufwies und eine konfirmatorische Untersuchung hier ebenfalls eine bessere Anpassung an die Daten ergab, fiel auch hier die Entscheidung zugunsten der 4-faktoriellen Lösung.⁵ Zur Benennung der Faktoren sind nur Subskalen berücksichtigt, deren Ladungen größer als .40 sind. Dies ist ein gängiges Kriterium, um die Bedeutsamkeit von Faktorladungen zu beurteilen (Bortz, 1999). Bei Schommer et al. (1992) verfehlten in der 4-faktoriellen Lösung die beiden Subskalen *Don't criticise authority* und *Knowledge is certain* mit Ladungen von .39 und .38 nur knapp dieses Kriterium, wurden jedoch aufgrund inhaltlicher Aspekte zur Interpretation des vierten Faktors herangezogen.

⁵ Schommer et al. (1992) nehmen Bezug auf eine dreifaktorielle Lösung aus der Studie von Schommer (1990). Leider ist in der zitierten Studie keine dreifaktorielle Lösung erwähnt. Es ist nicht zu bestimmen, um welche Publikation es sich handelt, da die entsprechende Referenz im Literaturverzeichnis nicht ausgewiesen ist.

Tabelle 4: Überblick über die Faktorenlösungen (Clarebout et al., 2001)

Untersuchung	Schommer	Schommer et al. (1992)		Schommer
	(1990)	3 Faktoren	4 Faktoren	(1993a)
N	266	424	424	n. a.
Analyse	Hauptfaktoren	Hauptfaktoren	Hauptfaktoren	Hauptfaktoren
Varianzaufklärung	55.2 %	46.2 %	54.3 %	keine Angabe
Rotation	Varimax	Varimax	Varimax	Varimax
Faktorname	F1: Innate Ability F2: Simple knowledge F3: Quick learning F4: Certain knowledge	F1: Innate Ability F2: Simple knowledge F3: Certain knowledge	F1: Externally controlled learning F2: Simple knowledge F3: Quick learning (F4: Certain Knowledge)	F1: Fixed ability F2: Simple Knowledge F3: Quick learning F4: Certain knowledge
Subskala				
Learning is quick	F3	F1	F3	F3
Can't learn how to learn	F1	F1	F1	F1
Learn first time	F1	F1	F1	F1
Concentrated effort is waste of time		F1	F1	
Success is unrelated to hard work	F1	F1	F1	F1
Avoid ambiguity	F2	F2	F2	F2
Seek single answers	F2	F2	F2	
Avoid integration		F2	F2	F2
Depend on authority		F2	F2	
Ability to learn is innate				F3
Don't criticise authority		F3	(F4, .39)	
Knowledge is certain	F4		(F4, .38)	F4

In den Ergebnissen der Arbeitsgruppe von Schommer zeigt sich somit keine empirische Evidenz für die postulierte Dimension *Structure of knowledge*. Dies wird auch durch direkte Replikationsstudien von Bendixen, Dunkle und Schraw (1994) sowie Paulsen und Wells (1998) gestützt. Es gibt allerdings weitere Hinweise auf eine Abweichung der empirisch gefundenen von der theoretisch postulierten Faktorenstruktur. Quian und Alvermann (1995) fanden mit

der von ihnen angepassten Version des EBQ Hinweise auf eine 3-faktorielle Struktur. Bezüglich der angenommenen und durch die Verwendung der Varimax-Rotation unterstellten Unabhängigkeit der Dimensionen epistemologischer Überzeugungen gibt es ebenfalls gegenteilige Befunde. So fanden Cole, Goetz und Willson (2000) Hinweise darauf, dass sich nur eine 2-faktorielle Struktur nachweisen lässt, deren beiden Faktoren sich als *nature of knowledge* und *nature of learning* identifizieren lassen. Dies deutet auf eine hierarchische Struktur der Dimensionen hin. Allerdings ist festzustellen, dass zwischen den Studien erhebliche methodische Unterschiede bestehen. So wurde z. B. nicht immer eine Faktorenanalyse auf Basis der Subskalen durchgeführt, sondern teilweise auf Ebene der Items vorgenommen. Es gibt allerdings auch Anzeichen für eine 5-faktorielle Struktur. So fanden z. B. Jehng, Johnson und Anderson (1993) mit einer abgewandelten Form des Fragebogens Hinweise auf fünf Dimensionen epistemologischer Überzeugungen. Ebenso fand sich bei der Analyse des auf dem EBQ beruhenden und aus 28 Items bestehenden *Epistemological Belief Inventory* (EBI) von Schraw, Bendixen und Dunkle (2002) eine 5-faktorielle Struktur, was auf die a priori postulierten fünf Dimensionen hindeutet. Das EBI ist darauf ausgelegt, die fünf Dimensionen im Sinne von Schommer (1990) zu erfassen. Allerdings wurde der EBI auf Basis der Items und nicht auf Basis vordefinierter Subskalen durchgeführt. Dies lässt erhebliche Zweifel an der gewählten Strategie der Nutzung von Subskalen aufkommen (vgl. Hall, Snell & Foust, 1999).

Ebenfalls direkte Replikationsstudien wurden mit dem ursprünglichen EBQ-Fragebogen von Clarebout et al. (2001) vorgenommen. In zwei Studien konnte die von Schommer (1990) postulierte Faktorenstruktur nicht nachgewiesen werden. DeBacker, Crowson, Beesley, Thoma und Hestevold (2008) versuchten ebenfalls eine Replikation der von Schommer postulierten Faktorenstruktur. Dazu wurde eine explorative Hauptachsen-Faktorenanalyse mit Varimax-Rotation auf der Basis der Subskalen durchgeführt und die Anzahl der zu extrahierenden Faktoren nach dem Kaiser-Kriterium bestimmt. Es ergab sich eine 3-faktorielle Lösung. Ebenso wurde eine Hauptachsen-Faktorenanalyse auf Basis der einzelnen Items des Schommer-Fragebogens durchgeführt. Das Kaiser-Kriterium ergab hier 22 zu extrahierende Faktoren. Auch hier ließ sich die postulierte Struktur nicht nachweisen. Festgehalten werden kann an dieser Stelle festhalten, dass Schommer (1990) trotz einiger Unstimmigkeiten in Bezug auf die postulierten fünf Dimensionen ein System „mehr oder weniger“ unabhängiger Dimensionen vorgeschlagen hat, das einige empirische Befunde bestätigen, viele andere hingegen nicht. Allerdings muss angemerkt werden, dass die gegenteiligen empirischen Befunde nicht gegen das von Schommer aufgestellte theoretische System epistemologischer Überzeugungen

an sich sprechen, sondern gegen das Instrument, mit dem dieses System quantitativ erfasst werden soll.

Wichtig für die vorliegende Arbeit ist die Schommers Ansatz zugrunde liegende inhaltliche Ausrichtung. Durch die Konzeptualisierung als dimensionales System und die Verwendung von Items zur Erfassung epistemologischer Überzeugungen sowie die Auswertung mithilfe faktorenanalytischer Verfahren wurde der Blickpunkt der Forschung auf den Bereich individueller Unterschiede gelegt. Instrumente wie dasjenige von Schommer gestatten es, die Unterschiede zwischen zwei oder mehreren Personen im Hinblick auf eine oder mehrere Variablen zu bestimmen. Im Sinne von Stern (1911) handelt es sich um eine differenzielle Sichtweise epistemologischer Überzeugungen und es liegt Variations- bzw. Korrelationsforschung vor. Damit wird die Forschung über epistemologische Überzeugungen aber auch in das Eigenschaftsparadigma der Persönlichkeitspsychologie gerückt (Asendorpf, 1999). Schommers *Epistemological-belief*-Modell ist daher der persönlichkeitspsychologischen Sichtweise epistemologischer Überzeugungen zuzurechnen. Dies impliziert aber auch eine statische Sichtweise epistemologischer Überzeugungen, wie sie dem Traitkonzept der Persönlichkeitspsychologie eigen ist, und rückt epistemologische Überzeugungen in die Position von Persönlichkeitseigenschaften. Eine weitere Folge ist, dass von individuellen Prozessen abgesehen und eine populationsbasierte Sichtweise eingenommen wird. Die Messung von Persönlichkeitseigenschaften im Rahmen von Traitkonzepten beinhaltet einen Vergleich der entsprechenden Eigenschaften der Mitglieder einer Referenzpopulation, d. h., die Ausprägung einer Eigenschaft wird als Vergleich mit der Ausprägung eines durchschnittlichen Mitglieds dieser Referenzpopulation angegeben (Asendorpf, 1999). Zusammenfassend lässt sich festhalten, dass sich epistemologische Überzeugungen im Rahmen persönlichkeitspsychologischer Ansätze und Methoden beschreiben lassen. Hierbei handelt es sich um ein System „mehr oder weniger“ unabhängiger Dimensionen, die sich als *Source of Knowledge*, *Certainty of knowledge*, *Structure of knowledge*, *Control of learning* und *Quick Learning* charakterisieren lassen.

Im Folgenden soll eine Erweiterung des Ansatzes von Schommer vorgestellt werden, der *Embedded-System-Ansatz* (EBS) von Schommer-Aikins (2004). Als Grundlage dienen die fünf Dimensionen von Schommer, die mit weitergehenden Überlegungen zur Verbindung von epistemologischen Überzeugungen zum Lernen und Lernerfolg, dem Einfluss der Kultur und dem Bedürfnis nach Balance in der Ausbildung epistemologischer Überzeugungen verknüpft

werden. In Abbildung 2 ist der *Embedded-System-Ansatz* grafisch dargestellt. Das Besondere des Embedded-System-Ansatzes ist einerseits, dass die epistemologischen Überzeugungen explizit in Überzeugungen zum Wissen, Wissenserwerb und Lernen aufgegliedert werden. Die Überzeugungen zum Wissen und Wissenserwerb beziehen sich auf die Struktur, die Quelle und die Rechtfertigung von Wissen, die Überzeugungen zum Lernen beziehen sich auf die Lerngeschwindigkeit und die Lernfähigkeit von Personen. Überzeugungen zum Lernen werden somit immer noch zum System epistemologischer Überzeugungen gezählt, ihre Wirkung auf andere Variablen wird jedoch gesondert dargestellt. Damit gibt Schommer die eingangs vorgestellte Position auf, wonach Überzeugungen zum Lernen den epistemologischen Überzeugungen gleichgesetzt werden. Zudem werden die epistemologischen Überzeugungen zum kulturellen Hintergrund einer Person sowie deren Leistung und deren selbstreguliertem Lernen in Beziehung gesetzt (s. Abbildung 2). Dabei stellen die kulturellen Einflüsse exogene Variablen des Modells dar, während die anderen Komponenten endogene Variablen sind.⁶

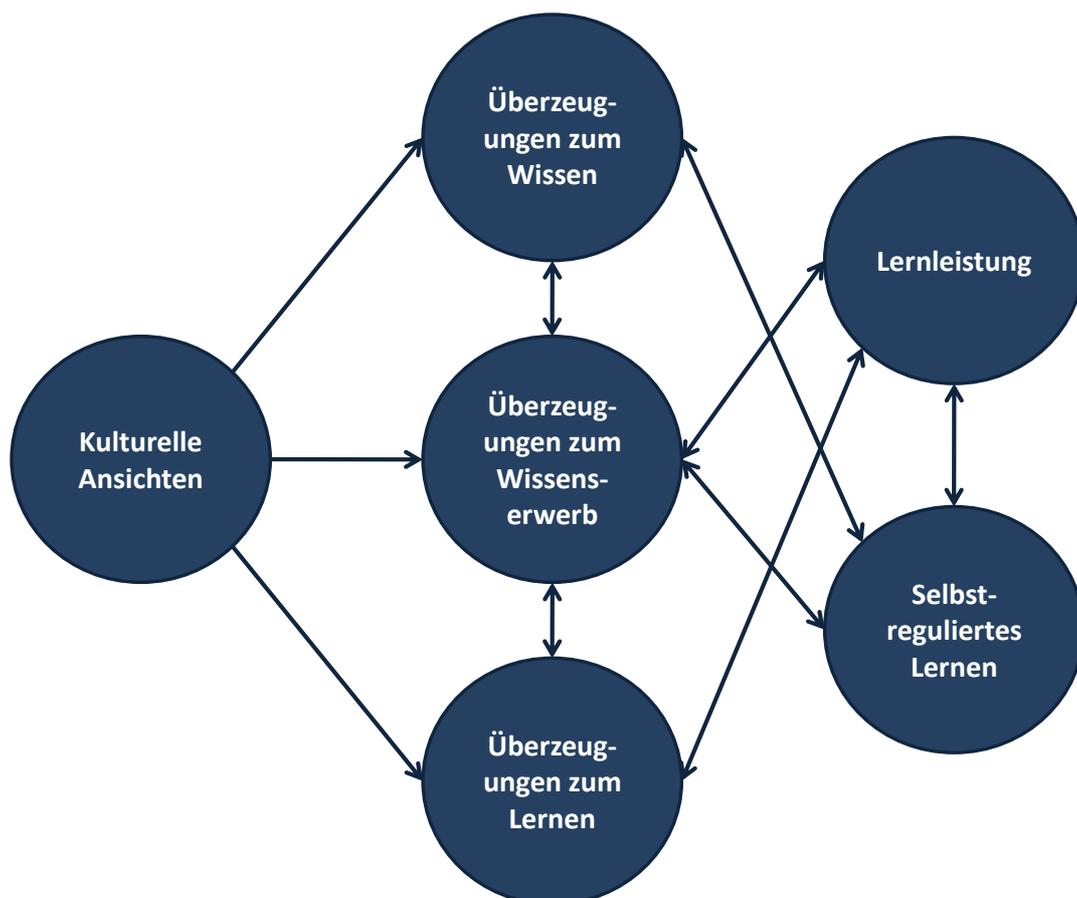


Abbildung 2: Embedded-Belief-System von Schommer-Aikins (2004)

⁶ In dieser Abbildung stellen exogene Variablen von außen an das Individuum herangetragene Größen dar, während endogene Variablen diejenigen sind, die im Individuum liegen.

Der kulturelle Hintergrund gibt die Kultur wieder, in welcher eine Person lebt. Dabei wird zwischen einer *kollektivistischen vs. individualistischen* Kultur und deren *horizontaler vs. vertikaler* Orientierung unterschieden (Triandis 1987, 1995). Die *kollektivistische vs. individualistische* Orientierung einer Kultur bestimmt, ob in dieser die Gruppe oder das Individuum im Mittelpunkt steht. In kollektivistischen Kulturen sind Individuen in einem engen sozialen Netz verbunden und Teil eines oder mehrerer Kollektive. Die Normen und Werte dieser Kollektive sind für das Individuum handlungsleitend und die Ziele der Kollektive haben Vorrang vor den Zielen des Individuums. In individualistischen Kulturen sieht sich das Individuum in erster Linie als Person und nicht als Teil eines Kollektivs. Handlungsleitend sind in erster Linie die Motive und Wünsche des Individuums und dessen persönlichen Ziele, die keinen Vorrang vor Zielen des Kollektivs haben. Die *horizontale vs. vertikale* Orientierung bezieht sich auf die Machtdistanz, d. h. darauf, ob es in einer Kultur ausgeprägte Hierarchien der Machtausübung gibt. In vertikal orientierten Kulturen sind ausgeprägte Machthierarchien zu finden, während diese in horizontal orientierten Kulturen nicht zu finden sind. Die *horizontale bzw. vertikale* Gliederung bestimmt somit, wie sich Personen mit fortgeschrittenen epistemologischen Überzeugungen gegenüber Autoritäten verhalten. Lerner in horizontal strukturierten Kulturen mit fortgeschrittenen epistemologischen Überzeugungen sehen sich in der Beziehung zu Experten auf gleicher Ebene, d. h., sie sehen sich mit dieser auf einer Stufe (Belenky et al., 1986, King & Kitchener, 1994). Baxter und Magolda (1992) konnten zeigen, dass die Neigung, Fragen zu stellen, abhängig von der Nähe zum Lehrenden ist, wobei in kollektivistischen Kulturen eine größere Nähe zu finden ist als in individualistisch geprägten Kulturen.

Der wesentliche Punkt des EBS besteht darin, dass die Zusammenhänge des kulturellen Einflusses mit epistemologischen Überzeugungen und die Einflüsse auf Lernleistung und selbst-reguliertes Lernen nicht auf theoretischer Ebene bleiben. Schommer-Aikins (2004) gibt eine Operationalisierung dieser Zusammenhänge im EBS an. Kulturelle Einflüsse und epistemologische Überzeugungen sowie deren Zusammenhang mit Überzeugungen zu Lernen, Lernerfolg, Lernstrategien, Verstehen, Metakognition, Problemlösen und Fähigkeiten zum kritischen Denken sind in Abbildung 3 dargestellt.

Das EBS ermöglicht es somit, die angenommenen Zusammenhänge einer Prüfung zu unterziehen, wobei sich diese Prüfung nicht ausschließlich auf den Zusammenhang von lediglich zwei Variablen beziehen muss. Die Konzeption des EBS legt eine Prüfung der Struktur des Zusammenhangs mehrerer Variablen nahe, wie dies im Rahmen von Strukturgleichungsmo-

dellen erfolgt. Weitere Möglichkeiten bietet die Berücksichtigung exogener Variablen (wie hier z. B. des kulturellen Einflusses). Im EBS wird diese Möglichkeit explizit in Betracht gezogen.

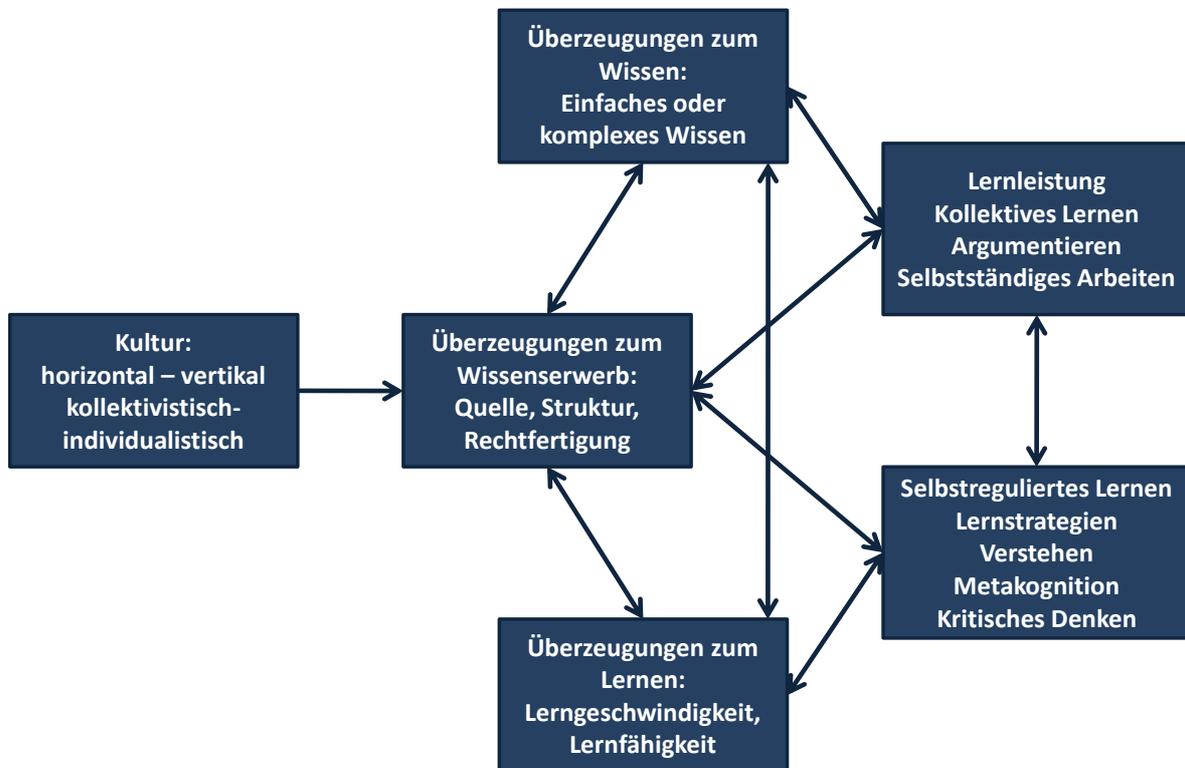


Abbildung 3: Angenommene Interaktion im EBS

1.3.2 Epistemologische Überzeugungen als Persönliche Theorien

Im letzten Kapitel wurde als Beispiel des persönlichkeitspsychologischen Ansatzes epistemologischer Überzeugungen das Modell von Schommer (1990) erläutert. Als weiterer Vertreter des persönlichkeitspsychologischen Ansatzes wird im Folgenden das Modell von Hofer und Pintrich (1997) beschrieben. Allerdings ist bei diesem Modell der Bezug zu interindividuellen Unterschieden nicht implizit in der Methodik verankert wie bei Schommer, sondern die Autoren konzipieren ihr Modell direkt als diejenigen Dimensionen, in denen sich Personen hinsichtlich ihrer epistemologischen Denkweisen unterscheiden können. Dieses Modell ist für die im Folgenden durchgeführten empirischen Untersuchungen nicht relevant, es wird aber wegen seiner Bedeutung für die Theoriebildung und die spätere Diskussion ausführlich präsentiert.

Die Konzeption epistemologischer Überzeugungen als Persönliche Theorie von Hofer und Pintrich (1997) resultiert aus einer Zusammenfassung der bis zu diesem Zeitpunkt wichtigsten

Modelle epistemologischer Überzeugungen. Dies sind in erster Linie die entwicklungspsychologischen Modelle sowie das Modell von Schommer. Das Anliegen von Hofer und Pintrich (1997) war es, die Bedeutung des Konstrukts durch eine eindeutige Definition hervorzuheben und die Dimensionen, welche den Kern epistemologischer Überzeugungen bilden, aufbauend auf dieser Definition festzulegen.

Laut Hofer und Pintrich (1997) wird eine eindeutige Definition epistemologischer Überzeugungen durch eine unterschiedliche Benennung sowie ein unterschiedliches Verständnis der zugrunde liegenden Konstrukte erschwert. Hofer und Pintrich (1997) teilen diese Ansätze in drei Klassen ein:

- Ansätze, die sich mit der kognitiven Entwicklung epistemologischer Überzeugungen befassen, worunter auch das hier entwicklungspsychologisch genannte Modell von Perry (1970) fällt,
- Ansätze, die epistemologische Überzeugungen als eine Menge von Einstellungen auffassen, die kognitive Prozesse beeinflussen wie das zuvor vorgestellte Modell von Schommer (1990), oder
- Ansätze, welche sich mit den kognitiven Prozessen selbst befassen.

Innerhalb der Klassen selbst gibt es weitere Abweichungen. So gehen die entwicklungspsychologischen Ansätze zwar von einer Abfolge von Stufen aus, unterscheiden sich aber in der Benennung sowie den Komponenten, welche die jeweiligen Stufen bilden. Ebenso unterscheiden sich die Ansätze, die epistemologische Überzeugungen als eine Menge von Einstellungen auffassen, in den Komponenten, die als epistemologische Überzeugungen angesehen werden (vgl. z. B. die Überzeugungen zum Lernen in der ursprünglichen Konzeption des Schommer'schen Modells).

Ein Vergleich der vorliegenden Modelle führte Hofer und Pintrich (1997) zu dem Schluss, dass, unabhängig von den speziell ausformulierten Komponenten, alle Modelle Überzeugungen zur Struktur des Wissens sowie Überzeugungen zum Prozess des Wissenserwerbs beinhalten. Durch die Beschränkung auf diese beiden Bereiche soll das Ziel einer eindeutigen Definition epistemologischer Überzeugungen und die Abgrenzung des Konstrukts gegenüber anderen Bereichen wie Lernen und kognitive Fähigkeiten erreicht werden. Diese beiden Bereiche passen in die Definition der Epistemologie als philosophische Disziplin (Hofer, 2001). Epistemologische Überzeugungen im Sinne des Modells von Hofer und Pintrich (1997) sind

demnach diejenigen Überzeugungen, die ein Individuum in Bezug auf die Natur des Wissens sowie den Prozess des Wissenserwerbs besitzt.

Neben der inhaltlichen Abgrenzung epistemologischer Überzeugungen stellt die Form epistemologischer Überzeugungen einen weiteren Punkt dar. Hofer und Pintrich (1997) postulieren epistemologische Überzeugungen als Persönliche Theorien. Damit sind epistemologische Überzeugungen in analoger Weise zu wissenschaftlichen Theorien organisiert. Diese Auffassung rückt epistemologische Überzeugungen in die Nähe subjektiver Theorien (Groeben, Wahl, Schlee & Scheele, 1988). Nach den Grundannahmen sind subjektive Theorien komplexe Zusammenhänge verschiedener Kognitionen, mit denen eine Person im Alltag versucht, ihre Welt zu strukturieren und Vorhersagen zu treffen. Demnach sind die subjektiven bzw. Persönlichen Theorien ähnlich strukturiert wie die Theorien, die von Wissenschaftlern zur Strukturierung von Daten und Vorhersage von Ereignissen benutzt werden. Nach Wellman (1990) müssen drei Kriterien erfüllt sein, um eine Menge von Annahmen als Persönliche Theorie zu bezeichnen.

Das erste Kriterium besagt, dass die in einer Theorie vorhandenen Annahmen und Konzepte untereinander kohärent sein müssen. Dies bedeutet, dass für das Vorliegen einer formalen Theorie die darin enthaltenen Konzepte vernetzt und widerspruchsfrei im Zusammenhang stehen müssen. Die epistemologischen Überzeugungen einer Person können somit von einfach nebeneinanderstehenden Konzepten und Annahmen bis zu einem komplexen Netz von Konzepten und Annahmen reichen. Nach Hofer und Pintrich (1997) zeigt die bisherige Forschung zu epistemologischen Überzeugungen, dass die Annahmen zu der Natur des Wissens sowie dem Prozess des Wissenserwerbs die Eigenschaften einer formalen Theorie aufweisen.⁷ Dies soll allerdings nicht bedeuten, dass die Persönlichen Theorien einer Person den gleichen Status erreichen wie die ausgearbeiteten philosophischen Erkenntnistheorien. Vielmehr können die epistemologischen Überzeugungen einer Person als eine Art naive Erkenntnistheorie gedeutet werden.

⁷ Über die genaue Art der Beziehungen, die den formalen Charakter einer Theorie bilden, machen Hofer und Pintrich (1997) allerdings keine Angaben. Nach der Standardtheorienkonzeption (vgl. Westermann, 2000) besteht eine wissenschaftliche Theorie aus einem axiomatisierten Ausgangssystem, das logisch formalisiert ist. Dazu gibt es eine Zweisprachenkonzeption sowie Zuordnungsregeln, welche die Begriffe der theoretischen Sprache denen der Beobachtungssprache zuordnen. Obwohl an persönliche Theorien nicht die gleichen Anforderungen gestellt werden wie an wissenschaftlichen Theorien, wäre es wünschenswert gewesen, dass diese Eigenschaften ausformuliert worden wären.

Das zweite Kriterium gesagt, dass zwischen den verschiedenen Entitäten und Prozessen eines Gegenstandsbereichs ontologische Unterschiede bestehen müssen. Dieses Kriterium gibt an, durch welche Entitäten ein Bereich – hier die epistemologischen Überzeugungen einer Person – gebildet werden und wie diese Entitäten untereinander strukturiert sind. Bei epistemologischen Überzeugungen lassen sich zwei grundlegend verschiedene Entitäten unterscheiden. Das sind zum einen Überzeugungen über die Natur des Wissens sowie zum anderen Überzeugungen zum Prozess des Wissenserwerbs. Diese Bereiche lassen sich weiter aufgliedern (s. u.) und bilden jeweils Einstellungen bezüglich zweier ontologisch verschiedener Entitäten ab. Gerber (2004) stellt die unterschiedliche Ontologie der Natur des Wissens und des Prozesses des Wissenserwerbs anschaulich dar, indem sie von „der Beschaffenheit des Wissens als Gegenstand (*nature of knowledge*) und der Beschaffenheit des Wissens als Prozess (*nature of knowing*)“ spricht.

Das dritte Kriterium sagt aus, dass eine Persönliche Theorie ein kausales und erklärendes Rahmenwerk für den jeweiligen Inhaltsbereich bieten muss. Die Relevanz epistemologischer Überzeugungen als kausales und erklärendes Rahmenwerk ist nach Hofer und Pintrich (1997) nicht so evident, wie dies für die naiven Theorien gilt, die für andere Bereiche wie z. B. physikalisches oder psychologisches Denken nachgewiesen sind (vgl. Sodian, 2008). Allerdings wird angenommen, dass Persönliche Theorien die Informationsverarbeitung beeinflussen und somit kausal relevant für das Verhalten einer Person sind und im gewissen Rahmen auch zur Erklärung herangezogen werden können.

Nach der Vorstellung des formalen Rahmens epistemologischer Überzeugungen als Persönliche Theorien ist noch der inhaltliche Bereich zu betrachten. Wie zuvor dargestellt, betrachten Hofer und Pintrich die Natur des Wissens sowie den Prozess des Wissenserwerbs als gemeinsamen Kern epistemologischer Überzeugungen, der durch Kriterien, die sich an dem Bereich der philosophischen Disziplin der Epistemologie orientieren, abgegrenzt ist. Diese beiden Bereiche lassen sich, wie in Tabelle 5 dargestellt, weiter aufgliedern.

Tabelle 5: Dimensionen epistemologischer Überzeugungen (Hofer und Pintrich, 1997, Gerber, 2004)

Natur des Wissens (<i>nature of knowledge</i>)		Prozess des Wissenserwerbs (<i>nature of knowing</i>)		
Englische Bezeichnung	Certainty of knowledge	Simplicity of knowledge	Source of knowledge	Justification of knowing
Deutsche Bezeichnung	Sicherheit des Wissens	Komplexität des Wissens	Wissensquelle	Wissensbegründung

Der Bereich *Natur des Wissens* gibt die Annahmen einer Person wieder, in welcher Art und Weise Wissen strukturiert ist. Dieser ist in den meisten Modellen epistemologischer Überzeugungen vorhanden, so z. B. auch in dem Modell von Schommer (1990). *Sicherheit des Wissens* gibt den Grad an, mit welchem eine Person Wissen als fest oder veränderbar ansieht. *Die Komplexität des Wissens* gibt den Grad an, mit welchem eine Person Wissen als eine Ansammlung distinkter Fakten oder als verbundene Konzepte betrachtet.

Der Prozess des Wissenserwerbs gliedert sich in die beiden Bereiche *Wissensquelle* sowie *Wissensbegründung*. *Wissensquelle* gibt an, woher eine Person ihr Wissen bezieht. Diese Dimension reicht von der Annahme, dass Wissen außerhalb der eigenen Person bei Autoritäten liegt, bis zu der Annahme, dass Wissen durch Reflexion in einen konstruktiven Prozess allein oder in Interaktion mit anderen gebildet werden kann. *Wissensbegründung* bezieht sich auf die Art und Weise, wie Personen Behauptungen evaluieren. Die Dimension bezieht sich auf den Gebrauch empirischer Evidenz sowie die Bewertung der Aussagen von Experten.

Alle Bereiche epistemologischer Überzeugungen weisen eine Entwicklungskomponente auf. So hat jede Dimension einen dualistischen und einen relativistischen Pol (Gerber, 2004), wobei dualistisch und relativistisch im Sinne von Perry (1970) zu verstehen sind. Nach Perry schreitet jedes Individuum von einem dualistischen Stadium zu einem relativistischen Stadium. Gerber (2004, S. 64) beschreibt diese Entwicklung wie folgt: „Eine dualistische Sichtweise beinhaltet eine kategorische Perspektive auf die Welt. Es gibt nur richtiges oder falsches Wissen (*'Sicherheit des Wissens'*), das aus einzelnen Fakten besteht (*'Komplexität des Wissens'*). Um dieses Wissen zu erlangen, muss auf Autoritäten zurückgegriffen werden (*'Wissensquelle'*). Darüber hinaus muss dieses Wissen nicht ständig neu überprüft oder hinterfragt werden (*'Wissensbegründung'*). Eine relativistische Sichtweise zeichnet sich demgegenüber durch die Annahme aus, dass Wissen sich immer wieder verändert und nicht ein für alle Mal feststeht (*'Sicherheit des Wissens'*). Wissen besteht aus einer Vielzahl miteinander verbundener Konstrukte (*'Komplexität des Wissens'*), die durchaus auch durch eigenes Nachden-

ken erfahrbar werden (*Wissensquelle*). Dieses Wissen sollte aber immer auch vor dem Hintergrund neuer Informationen evaluiert werden (*Wissensbegründung*).“

Die empirischen Anwendungen des Ansatzes von Hofer und Pintrich sind bislang nicht sehr zahlreich. Eine deutschsprachige Arbeit, die unter anderem eine Entwicklung eines Fragebogens zur Messung der vier Dimensionen beinhaltet, stammt von Gerber (2004). Weitere Anwendungen dieses Ansatzes stammen von Hofer (2000, 2004b) sowie Burr und Hofer (2002), wobei die Erfassung der Dimensionen nicht mit einem eigens entwickelten Instrument, sondern durch qualitative Interviews erfolgte. Im Rahmen der Anwendung von Modellen epistemologischer Überzeugungen in der Fachdidaktik einzelner Disziplinen wurden der Ansatz von Hofer und Pintrich von Elder (2002) sowie von Conley, Pintrich, Vekiri und Harrison (2004) zur Beschreibung und Messung epistemologischer Überzeugungen in den Naturwissenschaften angewandt. Eine deutschsprachige Variante des Instruments von Conley et al. (2004) findet sich in der Untersuchung von Urhahne und Hopf (2004).

In der schon genannten Arbeit von Gerber (2004) wurde ausgehend von Schommers Fragebogen (Schommer, 1990) ein deutschsprachiges Verfahren zur Erfassung epistemologischer Überzeugungen entwickelt, das die vier Dimensionen nach Hofer und Pintrich (1997) erfassen soll. Bei der Konstruktion stand die inhaltliche Validität der Skalen im Vordergrund, sodass für die Formulierung der Items ein facettentheoretisches Vorgehen gewählt wurde (vgl. Borg, 1992). Dazu wurden facettentheoretische Abbildungssätze aufgestellt, die zur Formulierung der Items dienten. Z. B. lautet der Abbildungssatz für die Dimension *Sicherheit des Wissens*:

$$\text{Person } p \text{ betrachtet die } \left\{ \begin{array}{l} \text{Erfahrbarkeit} \\ \text{Erklärbarkeit} \end{array} \right\} \text{ von Wissen } \left\{ \begin{array}{l} \text{bereichsübergreifend} \\ \text{bereichsspezifisch} \end{array} \right\} \text{ als } \left\{ \begin{array}{l} \text{gegeben} \\ \text{nicht gegeben} \end{array} \right\},$$

das heißt die Beschaffenheit von Wissen als $\left\{ \begin{array}{l} \text{sicher} \\ \text{relativ} \end{array} \right\}$.

Auf Basis dieser Abbildungssätze wurden Items entwickelt, welche die Sicherheit von Wissen domänenspezifisch (bereichsspezifisch) sowie domänunspezifisch (bereichsunspezifisch) erfassen sollten. Diese Art der Itementwicklung wurde auf alle vier Skalen angewandt. Insgesamt wurden für die vier Dimensionen 44 Items generiert, die anschließend einer Hauptkomponentenanalyse unterzogen wurden. Die Beantwortung erfolgte auf einer vierstufigen Ratingskala. Sowohl der Scree-Plot als auch die Ergebnisse von Parallelanalysen legten die Extraktion von vier Hauptkomponenten nahe. Nach der Extraktion von vier Hauptkomponenten mit anschließender Varimax-Rotation ergab sich eine Varianzaufklärung von 39%. Die Hauptkomponen-

ten ließen sich als *Wissensbegründung*, *Komplexität des Wissens*, *Wissensquelle – Autoritäten* sowie *Sicherheit des Wissens* interpretieren. Zur Konstruktion der jeweiligen Skalen wurden dann diejenigen Items ausgeschlossen, die nur eine geringe Ladung aufwiesen, bzw. nach Inspektion diejenigen Items hinzugenommen, die nicht theoriekonform luden. Tabelle 6 fasst die Ergebnisse der Skalenkonstruktion zusammen.

Tabelle 6: Übersicht über die Skalen von Gerber (2004)

Skala	Anzahl ursprünglicher Items	Anzahl Items nach Skalenkonstruktion	Cronbachs α
Sicherheit des Wissens	6	8	$\alpha = .72$
Komplexität des Wissens	17	10	$\alpha = .79$
Wissensbegründung	9	10	$\alpha = .87$
Wissensquelle - Autoritäten	12	7	$\alpha = .76$

Die Reliabilitäten sind bis auf eine Ausnahme nicht im Bereich größer als .80, der als zufriedenstellend gilt, aber im üblichen Bereich der Skalen, wie sie in Fragebogen zur Erfassung epistemologischer Überzeugungen vorkommen. Interessant sind die Korrelationen der Skalen untereinander. Zwar wurden die Hauptkomponenten orthogonal rotiert und somit a priori eine Unabhängigkeit der Dimensionen angenommen, diese Orthogonalität gilt allerdings nicht für die Summenwerte der Skalen.⁸ Die Interkorrelation der Skalen kann darüber Aufschluss geben, ob die gedachten Dimensionen epistemologischer Überzeugungen in realiter unabhängig sind oder nicht. Tabelle 7 fasst die Ergebnisse zusammen.⁹ Die absolute Höhe der Korrelationen insgesamt lässt darauf schließen, dass die postulierten Dimensionen epistemologischer Überzeugungen nicht unabhängig voneinander sind.

Tabelle 7: Interkorrelation der Skalen aus Gerber (2004)

	Sicherheit des Wissens	Komplexität des Wissens	Wissensbegründung	Wissensquelle – Autoritäten
Sicherheit des Wissens		.04	.09	.31**
Komplexität des Wissens			.59**	-.30**
Wissensbegründung				-.20**

Anmerkung: **: $p < .01$

⁸ Die Unabhängigkeit gilt nur für die Komponentenwerte der Hauptkomponenten. Durch Ausschluss und Hinzunahme von Items und bloßes Berechnen eines Summenwertes sind die errechneten Werte nicht mehr unabhängig voneinander.

⁹ Die Stichprobengröße wurde leider bei Gerber (2004) nicht angegeben.

Eine weitere Anwendung hat der Ansatz von Hofer und Pintrich in der Arbeit von Elder (2002) gefunden. Die Arbeit reiht sich in die Anwendung von Modellen epistemologischer Überzeugungen in der Didaktik der Naturwissenschaften ein. Aufbauend auf den vier Dimensionen *Sicherheit des Wissens*, *Komplexität des Wissens*, *Wissensquelle* und *Wissensbegründung* konzipierte Elder für die Naturwissenschaften die Dimensionen *Changing Nature of Science*, *Role of Experiments in Science*, *Coherence of Scientific Knowledge* und *Source of Scientific Knowledge*. *Changing Nature of Science* stellt die auf die Naturwissenschaften bezogene Dimension *Sicherheit des Wissens* dar. *Role of Experiments in Science* hat keine direkte Entsprechung in den Dimensionen von Hofer und Pintrich (1997). Der Inhalt dieser Dimension betrifft die Rolle von Experimenten zur Gewinnung naturwissenschaftlicher Erkenntnisse und bezieht sich somit auf die Quelle von Wissen. *Coherence of Scientific Knowledge* stellt eine Mischung der Struktur und der Quelle naturwissenschaftlichen Wissens dar. Die Pole werden durch einfach strukturiertes Faktenwissen gebildet, welches durch Experten vermittelt wird, gegenüber einem hoch komplexen Wissen aus verbundenen Konzepten, welches von einer Person erschlossen werden muss. *Source of Scientific Knowledge* ist die direkte Entsprechung zu der von Hofer und Pintrich (1997) beschriebenen Dimension *Source of Knowledge*.

Zu diesen Dimensionen entwarf Elder (2002) einen aus 25 Items bestehenden Fragebogen, die auf einer fünfstufigen Ratingskala von 211 Schülern einer Highschool beantwortet wurden. Die Items wurden a priori den vier Dimensionen epistemologischer Überzeugungen zugeordnet, analog dem Vorgehen von Schommer (1990). Items, welche negative Trennschärfen aufwiesen, wurden umkodiert. Zur Erhöhung der internen Konsistenz wurden einzelne Items anderen Skalen zugeordnet bzw. ausgeschlossen. Auf diese Weise entstanden drei Skalen: *Changing Nature of Science*, *Role of Experiments and Source of Knowledge (Reason)* und *Source of Knowledge (Authority)*.

Tabelle 8: Skalen zur Messung epistemologischer Überzeugungen nach Elder (2002)

Skala	Cronbachs α	Anzahl Items	Beschreibung
<i>Changing Nature of Science</i>	.67	9	Wissen verändert und entwickelt sich über die Zeit
<i>Role of Experiments and Source of Knowledge (Reason)</i>	.52	6	Wissen entsteht aus Tests und Experimenten und aus schlussfolgern und denken
<i>Source of Knowledge (Authority)</i>	.64	4	Wissen stammt von Autoritäten, einschließlich Büchern und Lehrern

Ausgehend von der Arbeit von Elder (2002) sowie dem Modell epistemologischer Überzeugungen als Persönliche Theorien von Hofer und Pintrich (1997) entwarfen Conley et al. (2004) einen 26 Items umfassenden Fragebogen. Der Fragebogen sollte die Dimensionen *Source*, *Certainty*, *Development* und *Justification* erfassen, er wurde von 187 Schülern zweimal im Abstand von neun Wochen bearbeitet. Tabelle 9 gibt einen Überblick über die psychometrischen Eigenschaften sowie die Definition der Dimensionen bei Conley et al. (2004). Die Items wurden wie bei dem Fragebogen von Elder (2002) den Skalen a priori zugeordnet und so gepolt, dass höhere Werte fortgeschrittenen¹⁰ epistemologischen Überzeugungen entsprechen.

Tabelle 9: Skalen zur Messung epistemologischer Überzeugungen (Conley et al., 2004)

Skala	Cronbachs α 1./2. Zeitpunkt	Anzahl Items	Beschreibung
<i>Source</i>	.81/.82	5	Wissen wird durch Autoritäten vermittelt
<i>Certainty</i>	.78/.79	6	Überzeugung in eine richtige Antwort
<i>Development</i>	.57/.66	6	Wissenschaft als sich entwickelndes und veränderndes Subjekt
<i>Justification</i>	.65/.76	9	Rolle von Experimenten und die Art, wie Individuen Wissen rechtfertigen

Zur Überprüfung der angenommenen Struktur epistemologischer Überzeugungen wurde zu beiden Zeitpunkten eine konfirmatorische Faktorenanalyse durchgeführt. Bei der CFA zur ersten Erhebung ergab sich zwar ein signifikanter χ^2 -Wert ($\chi^2 = 396.39$, $df = 293$, $\chi^2/df = 1.353$), der eine Abweichung des theoretisch angenommenen Modells von den empirischen Daten nahelegt. Die Fitindizes deuteten jedoch auf eine gute Anpassung der Daten an das Modell hin ($RMSEA = .038$, $CFI = .90$, $NNFI = .89$ und $RMR = .062$). Da sich keine Möglichkeiten zur Verbesserung des Modells ergaben, schlossen Conley et al. (2004) auf eine Gültigkeit des 4-faktoriellen Modells sowie der a priori angenommenen Itemzuordnung. Auch zeigte sich hier eine relativ starke Korrelation einzelner Dimensionen. Tabelle 10 gibt einen Überblick über die Korrelationen zum Zeitpunkt der ersten Erhebung. Vor allem die Korrelation der Skala zur Erfassung der Dimension *Source* mit der Skala zur Erfassung der Dimension *Certainty* sowie die Korrelation der Skala zur Erfassung von *Development* und *Justification* deuten auf eine starke Überlappung dieser Dimensionen hin.

¹⁰ Fortgeschritten bezieht sich auf den Ausprägungsgrad der Antworten auf den Ratingskalen, wobei gilt, dass höhere Skalenwerte als fortgeschritten angesehen werden. Dem liegt die normative Vorstellung zugrunde, dass sich epistemologische Überzeugungen in ein Kontinuum von naiv bis fortgeschritten einordnen lassen und fortgeschrittene epistemologische Überzeugungen elaborierte Vorstellungen zum Wissen abbilden und zu besseren Lernleistungen führen.

Tabelle 10: Interkorrelation der Skalen aus Conley et al. (2004)

	Certainty	Development	Justification
Source	.76**	.29**	.12
Certainty		.26**	.17*
Development			.47**

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

Eine Übersetzung des Fragebogens von Conley et al. (2004) in die deutsche Sprache wurde von Urhahne und Hopf (2004) vorgenommen. Dabei wurden die 26 Items mit Modifikationen zweier Items übernommen. Die Dimensionen wurden mit *Quelle des Wissens*, *Sicherheit des Wissens*, *Entwicklung des Wissens* und *Rechtfertigung des Wissens* bezeichnet. Die Anzahl der Items für jede Skala entspricht der Anzahl der Items bei Conley et al. (2004), die Polung erfolgte wiederum so, dass höhere Skalenwerte auf fortgeschrittene epistemologische Überzeugungen hinweisen. Auch hier erfolgte die Zuordnung der Items zu den Skalen a priori auf Grundlage der Zuordnung von Conley et al. (2004). Der Fragebogen wurde 167 Schülern zur Bearbeitung vorgelegt. Es ergaben sich für die vier Skalen *Quelle des Wissens*, *Sicherheit des Wissens*, *Entwicklung des Wissens* und *Rechtfertigung des Wissens* folgende interne Konsistenzen: .67, .41, .66 und .68 (Cronbachs α). Zur Überprüfung der faktoriellen Struktur führten Urhahne und Hopf (2004) eine Hauptkomponentenanalyse mit anschließender Varimax-Rotation und der Extraktion von vier Hauptkomponenten durch. Es ergab sich eine Varianzaufklärung von 39.8 %. Mit Ausnahme einiger Items mit bedeutenden Nebenladungen und einem Item der Skala *Sicherheit des Wissens* luden alle Items auf den ihnen zugeordneten Hauptkomponenten. Zusätzlich wurde eine konfirmatorische Faktorenanalyse durchgeführt. Wie bei Conley et al. (2004) deutete der signifikante χ^2 -Wert auf eine Abweichung der empirischen Daten vom theoretischen Modell hin ($\chi^2 = 521.21$, $df = 293$, $\chi^2/df = 1.78$), wobei auch hier die Fitindizes zufriedenstellende Werte erreichten ($RMSEA = .068$, $CFI = .98$ und $TLI = .98$). Es zeigten sich ebenfalls Korrelationen zwischen den Skalen, die in Tabelle 11 zu finden sind.

Tabelle 11: Korrelation der Skalen aus Urhahne und Hopf (2004)

	Quelle	Entwicklung	Sicherheit
Rechtfertigung	-.31**	-.37**	-.06
Quelle	-	.00	.31***
Entwicklung		-	.26**

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

Eine weitere und letzte Anwendung des Ansatzes von Hofer und (1997) stammt von Hofer (2000) selbst. Ihre Untersuchung zielte auf die Frage ab, ob es hinsichtlich unterschiedlicher akademischer Disziplinen Unterschiede in der Struktur epistemologischer Überzeugungen gibt. Hofer hat den von ihr entwickelten *Discipline-focused Epistemological Belief Questionnaire* (DEBQ) verwandt, um die vier Dimensionen des Modells von Hofer und Pintrich (1997) zu erfassen. Der DEBQ umfasst 27 Items, die aus Perrys (1970) *Checklist of Educational Values* sowie dem EBQ stammen bzw. eigens entwickelt wurden. Beide Fragebögen wurden von 326 Studenten auf einer fünfstufigen Ratingskala bearbeitet, wobei jeweils ein auf die Psychologie bzw. die Naturwissenschaften bezogener DEBQ vorgegeben wurde.

Die Auswertung erfolgte auch hier auf faktorenanalytischem Weg separat für jeden der beiden Fragebögen. Der Eigenwertverlauf einer explorativen Faktorenanalyse¹¹ legte bei beiden Faktorenanalysen die Extraktion von vier Faktoren nahe. Diejenigen Items, die auf diesen vier Faktoren luden, wurden für die nachfolgenden Analysen beibehalten. Anschließend wurden eine Hauptkomponenten- und eine Maximum-Likelihood-Faktorenanalyse mit Varimax-Rotation und der Vorgabe von vier Hauptkomponenten bzw. Faktoren durchgeführt. Die Ergebnisse der Hauptkomponentenanalyse entsprachen dem Ergebnis der Maximum-Likelihood-Faktorenanalyse, ein Item wies ein anderes Ladungsmuster auf. Es ergaben sich für jede Disziplin die folgenden vier Faktoren: *Certain/simple knowledge*, *Justification for knowing: personal*, *Source of Knowledge: authority* und *Attainability of truth*. Diese Faktoren klärten im DEBQ mit Ausrichtung auf die Psychologie 46.1 % und mit Fokus auf die Naturwissenschaften 53.1 % der Varianz auf. Die Items mit der Hauptladung auf dem jeweiligen Faktor wurden zur Bildung der Skalen herangezogen. Tabelle 12 gibt einen Überblick über die Eigenschaften der gebildeten Skalen. Der Faktor *Attainability of truth* bezieht sich auf die Frage, ob in den Wissenschaften letztendlich endgültige Antworten ermittelt werden können, und weist somit Bezüge zur Dimension *Sicherheit des Wissens* auf.

Tabelle 12: Interne Konsistenzen und Anzahl der Items der Skalen aus Hofer (2000)

Skala	Cronbachs α Psychologie	Cronbachs α Naturwissenschaften	Anzahl Items
<i>Certain/simple knowledge</i>	.74	.81	8
<i>Justification for knowing: personal</i>	.56	.61	4
<i>Source of Knowledge: authority</i>	.51	.64	4
<i>Attainability of truth</i>	.60	.75	2

¹¹ Hofer macht keine Angabe über die Art der Faktorenanalyse (PCA, PAF, Maximum Likelihood), welche den Eigenwertverläufen zugrunde liegt.

Obwohl hier keine Korrelation der Skalen vorgegeben wurde und auch die Extraktion orthogonal erfolgte, ergaben sich Überlappungen der Dimensionen. Die Items, welche ursprünglich für die Dimension *Sicherheit des Wissens (Certainty)* bzw. *Einfachheit des Wissens (Simplicity)* gedacht waren, luden auf einen Faktor. Dies deutet auf einen Zusammenhang dieser Dimensionen hin und lässt auf eine oblique Faktorstruktur schließen.

Das Modell von Hofer und Pintrich (1997) findet also empirische Unterstützung. Untermauert wird der Ansatz durch die Ergebnisse von Gerber (2004). Besonders hervorzuheben bei dieser Arbeit ist die facettentheoretische Konstruktion der Items, die theoretische Ausrichtung auf die vier von Hofer und Pintrich vorgeschlagenen Dimensionen epistemologischer Überzeugungen und die anschließende Skalenkonstruktion im Sinne der klassischen Testtheorie. Die Arbeiten von Conley et al. (2004) sowie Urhahne und Hopf (2004) lehnen sich eng an das Modell von Hofer und Pintrich an, sind aber in ihrer Ausrichtung speziell auf epistemologische Überzeugungen in den Naturwissenschaften bezogen. Besonders hervorzuheben ist das konfirmatorische Vorgehen in beiden Studien, auch wenn die von den Autoren berichteten Fitindizes nicht alle auf einen allgemeinen Modellfit hinwiesen. Die Arbeit von Elder (2002), welche den theoretischen Rahmen für Conley et al. (2004) sowie Urhahne und Hopf liefert, kann wegen der ausschließlichen Verwendung itemanalytischer Methoden nicht zur empirischen Bewertung des zugrunde liegenden theoretischen Modells herangezogen werden. Die Arbeit von Hofer (2000) liefert ebenfalls Belege für die vorgeschlagenen vier Dimensionen epistemologischer Überzeugungen. In allen vorgestellten Studien finden sich aber Hinweise, dass diese Dimensionen nicht unabhängig voneinander sind, und korrelieren die abgeleiteten Skalen teilweise stark miteinander. Es stellt sich also die Frage nach der Struktur epistemologischer Überzeugungen, die mithilfe obliquer oder hierarchischer Faktorenanalysen zu untersuchen bleibt.

Alle vorgestellten Untersuchungen verwendeten dieselbe Methodik und lassen sich der persönlichkeitspsychologischen Sichtweise epistemologischer Überzeugungen zuordnen. Die eigentliche Neuerung von Hofer und Pintrich (1997), die Konzeption epistemologischer Überzeugungen als Persönliche Theorien mit den daraus folgenden Eigenschaften, wird somit weitgehend nicht durch die bis dahin vorliegende Forschung abgedeckt. Vor allem die interne Struktur epistemologischer Überzeugungen wird durch die persönlichkeitspsychologische Methodik nur unzureichend erfasst. Die interne bzw. kognitive Struktur epistemologischer Überzeugungen wird im Ansatz epistemologischer Ressourcen von Hammer und Elby (2002)

thematisiert, der im übernächsten Kapitel vorgestellt wird. Zuvor sollen epistemologische Überzeugungen nach Moschner und Gruber (2003) dargestellt werden. Da sich ein Großteil der folgenden empirischen Fragestellungen auf den Fragebogen von Moschner und Gruber bezieht, erfolgt auch eine ausführlichere Darstellung.

1.3.3 Epistemologische Überzeugungen nach Moschner und Gruber

Im deutschsprachigen Raum haben Barbara Moschner und Hans Gruber einen eigenständigen Ansatz epistemologischer Überzeugungen entwickelt. Im Verständnis von Gruber und Moschner (2003) bestehen epistemologische Überzeugungen aus drei Komponenten:

- **Begabung:** wird als eine angeborene Fähigkeit beschrieben, welche allgemeine oder spezielle Leistungen von Individuen beeinflusst. Begabung wird dabei nur als indirekt erfassbar betrachtet. Lernen bedeutet eine stabile Veränderung von Verhalten sowie von kognitiven Prozessen und Strukturen.
- **Lernen:** wird in diesem Modell somit im Sinne behavioristischer und kognitionswissenschaftlicher Ansätze aufgefasst.
- **Wissen:** bezeichnet die kognitiven Strukturen, welche die Repräsentationen verschiedener Domänen enthalten. Dem Wissensbegriff liegt demnach eine kognitionswissenschaftliche Sichtweise zugrunde und Wissen wird als domänenspezifisch angesehen. Die Repräsentation einer Domäne umfasst auch Mechanismen zum Erwerb, zur Speicherung, zum Abruf und zur Nutzung von Wissensinhalten.

In den Alltagsvorstellungen der Lernenden und Lehrenden sind diese Komponenten verbunden und bilden die epistemologischen Überzeugungen einer Person. Epistemologische Überzeugungen sind nach Gruber und Moschner (2003) „intuitive und unbewusste Auffassungen von der Natur des Wissens, des Lernens und der Begabung, die einen Kontext bilden, auf dessen Grundlage eine Person auf intellektuelle Ressourcen zugreift bzw. diese nutzbar macht.“ Diese epistemologischen Überzeugungen sind direkt relevant für Lehren und Lernen. Gruber und Moschner zufolge haben die epistemologischen Überzeugungen von Lernenden Einfluss auf den Lernerfolg. Ebenso wirken die epistemologischen Überzeugungen von Lehrenden auf die Gestaltung von Lernangeboten, die Unterstützung bei der Selektion geeigneter Lernstrategien von Lernenden und die Förderung von Begabung und Wissenserwerb. Dabei spielt aber auch der jeweilige fachwissenschaftliche Erkenntnisstand eine Rolle.

Die Zusammenhänge von epistemologischen Überzeugungen, fachwissenschaftlichem Erkenntnisstand und der didaktischen Gestaltung von Lehr-Lern-Situationen sind in Abbildung 4 dargestellt.

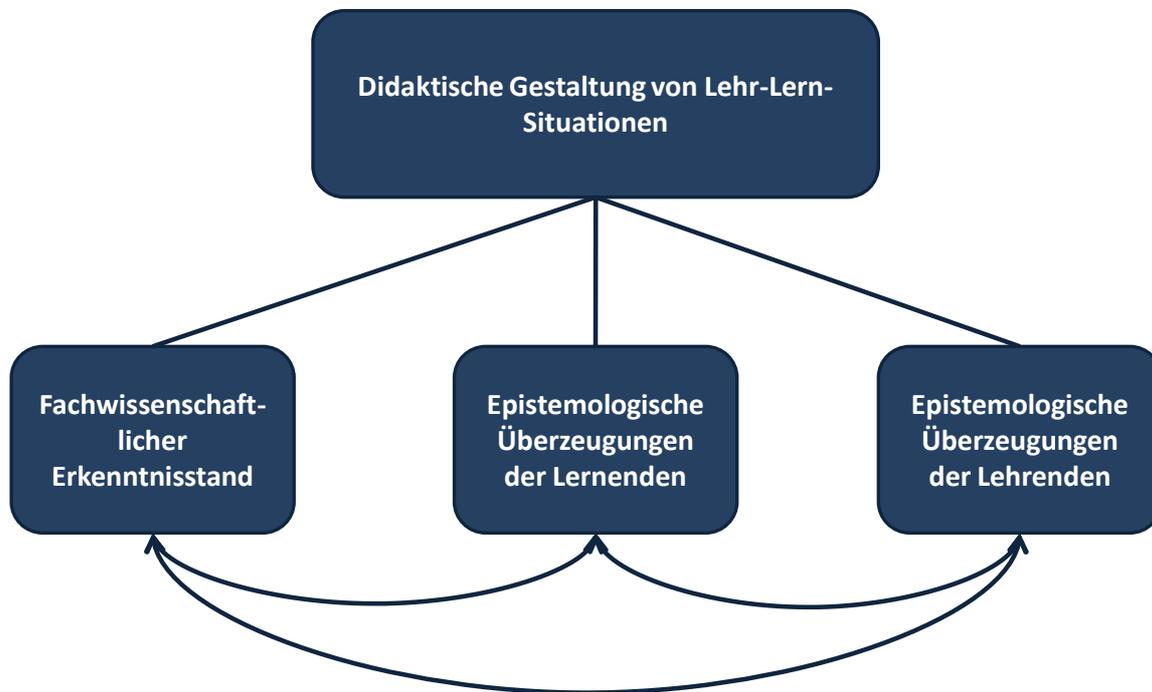


Abbildung 4: Rolle epistemologischer Überzeugungen in der Gestaltung von Lehr-Lern-Situationen (Gruber & Moschner, 2003)

Epistemologische Überzeugungen sind somit ein wichtiger Mediator zwischen dem jeweiligen Wissen in einem Fach, der didaktischen Aufbereitung durch Lehrende und dem Lernerfolg bei Lernenden.

Nach erfolgter Abgrenzung der drei grundlegenden Komponenten epistemologischer Überzeugungen und der Beschreibung des Zusammenhangs zwischen Lernerfolg und didaktischer Gestaltung, bleibt noch offen, aus welchen Dimensionen sich die Komponenten zusammensetzen. Dabei verstehen Moschner und Gruber (2005a) epistemologische Überzeugungen als ein breites Konstrukt, welches auch Überzeugungen über geschlechts- und kulturspezifische Aspekte des Wissens, der Begabung und des Lernens enthält. Zu dieser Konzeption wurde ein Fragebogen entwickelt, welcher die quantitative Erfassung epistemologischer Überzeugungen ermöglicht. Dieser Fragebogen basiert auf einer Durchsicht der einschlägigen Literatur auf dem verbreiteten Fragebogen von Schommer (1990) und erweitert diesen um die Konstrukte, die nicht in der Konzeption von Schommer enthalten sind. Moschner et al. (2005) haben a priori folgende die in Tabelle 13 für jede der Komponenten Wissen, Begabung und Lernen aufgeführten Konstrukte angenommen, für die dann verschiedene Items zur Erfassung gene-

riert wurden. Im Folgenden werden diese Konstrukte in Anlehnung an Moschner et al. (2005) näher beschrieben und ihre Bedeutung und Herkunft erläutert.

Tabelle 13: A priori Konstrukte nach Moschner und Gruber (2005)

1.	Wissen
1.1	Umgang mit Autoritäten (Source of Knowledge)
1.2	Objektivität des Wissens (Certain Knowledge)
1.3	Soziale Komponente des Wissens
1.4	Value of Knowledge
1.5	Reflexive Natur des Wissens
1.6	Silence of Knowledge
1.7	Fachabhängigkeit von Alltagswissen
2.	Lernen
2.1	Strategische Komponente des Lernens ¹²
2.2	Quick Learning
2.3	Geschlechtsspezifische Wissenszugänge
2.4	Kulturspezifische Wissenszugänge
3.	Begabung
3.1	Fixed Ability

Das Konstrukt *Umgang mit Autoritäten* bezieht sich auf die Rolle von Autoritäten bei der Wissensvermittlung. Autoritätsgläubige Menschen nehmen dabei an, dass Personen, die bereits gründliches und durch die Scientific Community bestätigtes Wissen in einem Themengebiet erworben haben, eine besondere Machtstellung bei der Vermittlung von Wissen zukommt. Das Konstrukt *Umgang mit Autoritäten* wird durch zwei Extreme beschrieben, die zugleich Auskunft über die Qualität des vermittelten Inhalts geben: Auf der einen Seite wird durch die Autoritäten vermitteltes Wissen akzeptiert, ohne von den Lernenden hinterfragt zu werden. Auf der anderen Seite steht die relativistische Annahme, dass auch eine Autorität und das vermittelte Wissen hinterfragt werden dürfen.

Objektivität des Wissens zielt auf die Frage, ob Wissen als objektiv angenommen werden kann und ist eine Grundfrage in der Beurteilung der Rolle, die Wissenschaft für die Welt spielt. Sie spiegelt insbesondere den Umgang mit dem von der Wissenschaft propagierten Wissen wider, was durch die beiden Pole „Wissen ist unumstößlich und ändert sich nicht“ und „Wissen ist vorläufig, entwickelt sich ständig weiter und muss daher hinterfragt werden“ ausgedrückt wird. Die Überzeugungen über die Objektivität von Wissen fallen in unterschied-

¹² Moschner et al. (2005) bezeichnen dieses Konstrukt als *Lernen lernen*.

lichen Domänen unterschiedlich aus; beispielsweise wird in der Mathematik eher von der Sicherheit von Wissen ausgegangen als in den Sozialwissenschaften (Schommer & Walker, 1995).

Das Konstrukt *Soziale Komponente des Wissens* wurde von Moschner et al. (2005) in Anlehnung an konstruktivistische Vorstellungen entworfen, die das Ausmaß der sozialen Komponente des Wissens thematisieren. Dies kann im Extremfall bedeuten, dass nur solches Wissen als gültig angesehen wird, welches mit anderen geteilt wird, also beispielsweise in der Literatur publiziert wurde und von der Scientific Community anerkannt ist.

Ebenso wie bei dem Konstrukt *Soziale Komponente des Wissens* handelt es sich bei *Value of Knowledge* um ein von Moschner et al. (2005) neu entwickeltes Konstrukt, welches auf die Frage abzielt, inwieweit Individuen die Bedeutung von Wissen an dessen gesellschaftlichem Nutzen messen. Die Einstellungen, der Wert von Wissen sei ausschließlich bzw. gar nicht von Relevanz für die Gesellschaft, werden als Pole des Konstrukts angesehen.

Ausgehend von dem Ansatz von King und Kitchener (1994), wonach ein Individuum unterschiedliche Stufen der Reflexion über Wissen durchläuft, wurde das Konstrukt *Reflexive Natur des Wissens* eingeführt. Darunter wird die Veränderlichkeit von Wissen in Abhängigkeit von der Perspektive und Erfahrung des Lernenden verstanden. Die beiden Pole werden durch die beiden Annahmen repräsentiert, dass Wissen entweder als absolut oder als relativ zur Perspektive und Erfahrung des Lernenden gesehen wird.

Das Konstrukt *Silence of Knowledge* findet seinen Ausgangspunkt in der von Belenky et al. (1986) gestellten Frage: „How can so many intelligent women feel so dull?“ *Silence of knowledge* bezeichnet eine Position, in der sich ein Individuum nicht in der Lage fühlt, sich auszudrücken, unabhängig davon, ob es viel weiß. Diese Sprachlosigkeit resultiert aus dem Gefühl der Unterlegenheit – oft von Frauen gegenüber Männern – und der Angst vor den Konsequenzen einer Äußerung. *Silence of Knowledge* kann sowohl situationsspezifisch (eine Person wird „zum Schweigen gebracht“) als auch situationsübergreifend (eine Person sagt generell eher wenig) sein.

Die *Fachabhängigkeit von Alltagswissen* kennzeichnet den Zusammenhang von Fach- und Allgemeinwissen. Das Konstrukt zielt auf die Vorstellung ab, dass Fachkenntnisse das All-

gemeinwissen beeinflussen. Dies impliziert weiterhin, dass sich allgemeine epistemologische Überzeugungen bei Vertretern unterschiedlicher Fächer unterschiedlich darstellen, da die allgemeinen epistemologischen Überzeugungen von den fachspezifischen epistemologischen Überzeugungen beeinflusst werden sollen.

Das Konstrukt *Strategische Komponente des Lernens* beschreibt die strategischen Komponenten des Lernens und kann damit als Gegenstück zum Konstrukt *Fixed Ability* gesehen werden. Auch dieses Konstrukt wurde von Moschner et al. (2005) neu entwickelt. Es beschreibt den Einfluss von Lerntechniken auf den Erfolg beim Wissenserwerb. Das Konstrukt wird durch die beiden Pole beschrieben, die besagen, dass Lernstrategien einen starken Einfluss bzw. keinen Einfluss auf den Erfolg beim Wissenserwerb haben.

Quick Learning beschreibt in Anlehnung an Schommer (1990) die Überzeugung, dass der Prozess des Wissenserwerbs schnell und erfolgreich oder gar nicht abläuft. Ein längerfristiges Arbeiten an einem Problem wird von einer Person als nicht erfolgversprechend angesehen.

Geschlechtsspezifische Wissenszugänge will die Überzeugungen bezüglich des Einflusses des Geschlechts auf das Wissen und den Wissenserwerb erfassen. Ausgehend von den Modellen von Belenky et al. (1986) und Baxter Magolda (1992) wurden Items konstruiert, welche die Einstellungen der Versuchspersonen zum Zusammenhang von Geschlecht und Wissen bzw. Wissenserwerb erfassen sollten.

Das Konstrukt *Kulturspezifische Wissenszugänge* erfasst den Einfluss, den die Kultur auf die Überzeugungen über Wissen und die Überzeugungen über den Prozess des Wissenserwerbs hat.

Das Konstrukt *Fixed Ability* geht auf Schommer (1990) in Anlehnung an Dweck und Leggett (1988) zurück und beschreibt die Annahmen, ob Lernfähigkeit angeboren ist oder sich im Laufe des Lebens verändern kann.

Zur Erfassung dieser Konstrukte wurde ein Itempool zusammengestellt. Dabei wurden von Moschner et al. (2005) anfänglich 98 Items generiert, welche die zuvor vorgestellten Konstrukte erfassen sollten, wobei für *Quick Learning* keine Items erstellt wurden und diese Dimension auch nicht weiter berücksichtigt wurde. Diese Items wurden von 114 Versuchspersonen

sonen, die an der Sommerakademie der Studienstiftung des deutschen Volkes 2005 teilnahmen, bearbeitet. Die Items wurden auf einer sechsstufigen Ratingskala, die von „trifft überhaupt nicht zu“ bis „trifft voll und ganz zu“ reicht, vorgelegt. Die Versuchspersonen teilten sich in 59 Frauen und 55 Männer mit einem mittleren Alter von 25.7 Jahren ($s = 2.8$ Jahre) auf. Da der Stichprobenumfang im Verhältnis zur Zahl der Items begrenzt war, wurden im ersten Schritt die den jeweiligen Konstrukten zugeordneten Items einer Itemanalyse unterzogen. Es wurden dann diejenigen Items entfernt, die eine Trennschärfe kleiner als .30 aufwiesen. Der auf diese Weise entstandene und reduzierte Itempool wurde einer Hauptkomponentenanalyse mit Varimax-Rotation unterzogen. Items mit einer Hauptladung größer als .40 und Nebenladungen kleiner als .40 wurden zu einer Skala zusammengefasst. Für diese neu gebildeten Skalen wurde Cronbachs α als Maß der internen Konsistenz berechnet. Items, welche die interne Konsistenz verminderten, wurden aus der Skala ausgeschlossen. Ebenso wurden Skalen, deren Reliabilität unter dem Wert von .60 lag, ausgeschlossen. Auf diese Weise wurden insgesamt neun Skalen gebildet. Die Konstrukte *Sicherheit des Wissens* und *Fachspezifische Wissenszugänge* konnten nicht reliabel erfasst werden. Eine Übersicht über die Skalen ist in Tabelle 14 zu finden.

Tabelle 14: Ergebnisse aus Moschner et al. (2005)

Skala	Anzahl der Items	Cronbachs α	M	SD
Umgang mit Autoritäten	6	.83	1.0	.4
Soziale Komponente des Wissens	9	.79	2.0	.5
Value of Knowledge	8	.74	1.8	.9
Strategische Komponente des Wissens	6	.67	3.2	.5
Reflexive Natur des Wissens	7	.62	4.0	.2
Geschlechtsspezifische Wissenszugänge	10	.91	2.1	.5
Kulturspezifische Wissenszugänge	7	.81	3.5	.3
Silence of knowledge	6	.70	3.3	1.0
Fixed Ability	5	.64	2.6	.8

Ausgehend von dieser Studie wurden weitere Studien zur Validierung dieses Instruments durchgeführt (Moschner & Gruber, 2005a, 2005b). In einem ersten Schritt wurden Items hinzugefügt, sodass ein Itempool von insgesamt 142 Items zur Verfügung stand, die den oben beschriebenen zwölf Konstrukten zugeordnet waren. Diese Items wurden von 508 Versuchspersonen auf einer sechsstufigen Ratingskala, die von „trifft überhaupt nicht zu“ bis „trifft voll und ganz zu“ reicht, bearbeitet. Die Erhebung war als Onlineuntersuchung ausgelegt, die in der zweiten Jahreshälfte bis Ende 2004 angeboten wurde. Von den 508 Versuchspersonen

waren 365 Frauen und 143 Männer, das mittlere Alter betrug 28.8 Jahre ($s = 4.2$). Die Verteilung der Semesterzahl hatte den Modus 3 und den Median bei 5 Semestern.

Die gesammelten Daten wurden einer explorativen Faktorenanalyse und einer anschließenden Itemanalyse unterzogen.¹³ Bei der Auswertung ergab sich, dass die Dimensionen *Umgang mit Autoritäten* und *Objektivität des Wissens* zusammenfallen. Die neue Dimension wird im Folgenden *Absolutes Wissen* bezeichnet. Ebenso fielen die Dimensionen *Fixed Ability* und *Strategische Komponente des Lernens* zusammen, diese Dimension wird im Folgenden als *Lernen lernen* bezeichnet. Die Dimensionen *Silence of Knowledge* und *Soziale Komponente des Wissens* fielen ebenso zusammen und werden im Folgenden als *Soziale Komponente des Wissens* bezeichnet. Die Skalen, welche reliabel erfasst werden konnten, sind in Tabelle 15 dargestellt.

Tabelle 15: Ergebnisse aus Moschner und Gruber (2005a)

Skala	Anzahl der Items	Cronbachs α
Absolutes Wissen	15	.82
Value of Knowledge	5	.76
Lernen lernen	6	.71
Reflexivität von Wissen	5	.77
Geschlechtsspezifische Wissenszugänge	10	.88
Kulturspezifische Wissenszugänge	7	.81
Soziale Komponente des Wissens	9	.75

Zur Validierung dieser Ergebnisse wurde eine erneute Onlineuntersuchung durchgeführt. An dieser nahmen insgesamt 116 Versuchspersonen teil, davon 90 Frauen und 26 Männer. Das mittlere Alter betrug 25.2 Jahre ($s = 4.2$). Die Verteilung der Semesterzahl hatte sowohl den Modus als auch den Median bei 6 Semestern. Die Untersuchung wurde im Sommer 2005 im Internet angeboten. Insgesamt wurden sieben Dimensionen mit insgesamt 53 Items auf einer sechsstufigen Ratingskala, die von „trifft überhaupt nicht zu“ bis „trifft voll und ganz zu“ reicht, bearbeitet. In Tabelle 16 sind die Ergebnisse aufgeführt.

¹³ Leider liegen die genauen Beschreibungen der durchgeführten statistischen Prozeduren nicht vor, die im Text genannten Studien sind in nicht publizierten Konferenzbeiträgen beschrieben.

Tabelle 16: Ergebnisse aus Moschner und Gruber (2005a)

Skala	Anzahl der Items	Cronbachs α
Reflexivität des Wissens	k. A.	k. A.
Absolutes Wissen	12	.82
Value of Knowledge	5	.79
Lernen lernen	6	.76
Geschlechtsspezifische Wissenszugänge	10	.86
Kulturspezifische Wissenszugänge	7	.87
Soziale Komponente des Wissens	5	.74

Anmerkung: k. A.: keine Angabe

Aus der Ergebnisdarstellung geht nicht hervor, welche Anzahl von Items und welche Reliabilität die Skala *Reflexive Natur des Wissens* hat. Auch geht nicht daraus hervor, in welcher Weise die Skalen mit den insgesamt 53 Items für die Berechnung der Reliabilität noch einmal verändert wurden.¹⁴ Für alle Skalen sind die internen Konsistenzen größer als .70. Der Fragebogen, der für die zweite Onlinestudie benutzt wurde und welcher auch in der vorliegenden Arbeit benutzt wird, ist in Anhang B zu finden.

Moschner und Gruber entwickelten das Instrument zum *Fragebogen zur Erfassung epistemologischer Überzeugungen* (FEE) weiter (Moschner & Gruber, im Druck), allerdings liegen zu dieser Weiterentwicklung keine publizierten Daten vor. Eine Übersicht über die Skalen mit der Anzahl der Items und den internen Konsistenzen ist in Tabelle 17 zu finden, der komplette FEE ist in Anhang C dargestellt.

Tabelle 17: Skalen des FEE

Skala	Anzahl der Items	Cronbachs α
Sicherheit von Wissen	7	.78
Umgang mit Autoritäten	6	.71
Reflexive Natur von Wissen	5	.79
Soziale Komponente von Wissen	5	.65
Wertigkeit von Wissen	5	.72
Geschlechtsspezifische Wissenszugänge	5	.83
Kulturspezifische Wissenszugänge	6	.78
Lernen lernen	4	.73

Auffällig beim FEE ist, dass der Faktor *Absolutes Wissen* in die beiden Faktoren *Sicherheit von Wissen* und *Umgang mit Autoritäten* aufgeteilt wurde (vgl. Anhang B und C). Obwohl die Konstruktion des FEE unbekannt ist, bietet die finale Fassung dieses Fragebogens die Möglichkeit, die Ergebnisse der durchgeführten weiterführenden Untersuchungen des EPI-

¹⁴ Aus den Ergebnisdarstellungen wird nicht ersichtlich, welche Items im Zuge der Skalenkonstruktion von der Vorstudie über die beiden Onlinestudien den jeweiligen Skalen zugerechnet wurden bzw. welche Items zwischen den Skalen verschoben oder auch entfernt wurden.

Fragebogens mit Skalen des FEE zu vergleichen. Somit können Angaben zu konvergierenden oder divergierenden Befunden gemacht werden.

Der Ansatz von Moschner und Gruber ist wie der von Schommer (1990) sowie Hofer und Pintrich (1997) im Bereich der Persönlichkeitspsychologie angesiedelt, d. h., epistemologische Überzeugungen werden als Traits und damit als stabile Eigenschaften einer Person gesehen. Da der Ansatz von Moschner und Gruber noch relativ neu ist, existieren dazu keine weiteren empirischen Untersuchungen. Insbesondere gibt es keine Untersuchungen zur Validität des Messinstruments und auch keine Untersuchungen, die das Instrument in der pädagogischen Praxis zur Messung von epistemologischen Überzeugungen angewandt haben. In der vorliegenden Arbeit wird dieses Instrument neben dem EBQ von Schommer als weiterer Schwerpunkt betrachtet.

Bis hier wurden detailliert drei Modelle epistemologischer Überzeugungen vorgestellt, die alle der persönlichkeitspsychologischen Perspektive zuzurechnen sind. Persönlichkeitspsychologische Modelle stellen aktuell einen Großteil der Studien zu epistemologischen Überzeugungen dar. Aus diesem Grund wird sich auf diese Arbeiten bezogen, wobei das Modell von Schommer und das von Moschner und Gruber mit dem dazugehörigen Fragebögen im Mittelpunkt stehen werden. Es existieren jedoch noch weitere Modelle epistemologischer Überzeugungen, die sich teilweise ausdrücklich von dem auf Traitkonzepten bezogenen persönlichkeitspsychologischen Ansatz distanzieren. Diese besitzen eine gänzlich andere theoretische Grundlage oder betrachten besondere Aspekte epistemologischer Überzeugungen. Insbesondere zu nennen sind hier kognitive Aspekte. Im Folgenden werden die bekanntesten dieser Ansätze vorgestellt.

1.4 Epistemologische Ressourcen

Als erstes dieser alternativen Modelle epistemologischer Überzeugungen wird der Ansatz epistemologischer Überzeugungen nach Hammer und Elby (2002) vorgestellt. Das Modell von Hammer und Elby verlässt den persönlichkeitspsychologischen Rahmen und betont ontologische Fragen sowie die kognitive Architektur epistemologischer Überzeugungen. Durch den Bezug zur Physikdidaktik ist dieses Modell in die Forschung über epistemologische Überzeugung innerhalb der Fachdidaktik einzelner Disziplinen einzuordnen.

Der Ansatz epistemologischer Ressourcen (Hammer & Elby, 2002, Louca, Elby, Hammer & Kagey, 2004) greift die Frage nach der Ontologie epistemologischer Überzeugungen auf und es wird weiter nach der internen Struktur einer informellen (persönlichen) Epistemologie gefragt. Damit einhergehend wird auch die Ebene aggregierter psychologischer Konstrukte wie Persönlichkeitsdimensionen oder Persönliche Theorien verlassen. Nach Elby und Hammer (2002) wird dabei implizit angenommen, dass Traits bzw. Persönliche Theorien unitäre psychologische Einheiten sind, die sich nicht weiter untergliedern lassen. Diese Unitaritätsannahme gilt in zwei Varianten: für die Vorstellung epistemologischer Überzeugungen als Persönliche Theorien sowie für die Vorstellung epistemologischer Überzeugungen als Traits:

- Persönliche Theorien sehen epistemologische Überzeugungen als deklaratives Wissen, zu dem es einen bewussten Zugang gibt.
- Traits sehen epistemologische Überzeugungen als dispositionelle Eigenschaften.

Beide Ansätze verwenden zur Erfassung epistemologischer Überzeugungen Fragebogenverfahren. Die zugrunde liegende Unitaritätsannahme lässt sich allerdings in beiden Ansätzen nicht rechtfertigen. So setzt der Ansatz epistemologischer Überzeugungen als Persönliche Theorien einen direkten Zugang zum deklarativen Wissen voraus. Deklaratives Wissen stellt Sachwissen dar (Edelmann, 1996), zu dem es einen direkten und bewussten Zugang gibt. Versuchspersonen sind nach diesem Ansatz also in der Lage, bei der Beantwortung eines Fragebogenitems direkt bzw. nach einigem Nachdenken auf das relevante Wissen zuzugreifen. Elby und Hammer halten diese Annahme für nicht haltbar, da epistemologische Überzeugungen nicht direkt abrufbar sind und erst in einem Prozess bewusster Reflexion zugänglich werden. Betrachtet man epistemologische Überzeugungen als Traits, so sind die epistemologischen Überzeugungen zwar nicht bewusst, können aber durch Präferenzen, Verhalten, Gewohnheiten oder andere Indikatoren erschlossen werden.¹⁵

In beiden Ansätzen wird eine Konsistenz epistemologischer Überzeugungen vorausgesetzt. Dies ist eine Folgerung aus der Annahme unitärer psychologischer Konstrukte. Für Elby und Hammer ist die Annahme einer Variation epistemologischer Überzeugungen in verschiedenen Kontexten jedoch naheliegender als die Annahme der Konstanz. Konsistenz epistemologischer Überzeugungen zwischen verschiedenen Kontexten kann Elby und Hammer zufolge

¹⁵ Elby und Hammer verwenden zur Verdeutlichung dieses Sachverhalts das anschauliche Beispiel eines Golfers beim Abschlag. Im Sinne des Ansatzes Persönlicher Theorien würde man einen Golfer danach fragen, ob er beim Abschlag aus- oder einatmet. Es wird angenommen, dass der Golfer dies weiß und darüber berichten kann. Im Sinne des Traitansatzes würde man das Verhalten eines Golfers während des Spiels durch das Verhalten während anderer Aktivitäten erschließen, z. B. wird ein Rechtshänder auch den Golfschläger mit der rechten Hand führen oder man würde den Golfer mithilfe eines Fragebogens nach seiner Händigkeit befragen.

auftreten, ist jedoch nicht zwingend. Verschiedene Studien zur Domänenabhängigkeit epistemologischer Überzeugungen legen nahe, dass diese zwischen verschiedenen Domänen variieren (wobei hier Domäne als ein spezieller Kontext ausgelegt wird). So fand Hofer (2000) Unterschiede in der Ausprägung epistemologischer Überzeugungen zwischen Studenten der Psychologie und der Naturwissenschaften. Ebenso fanden Stodolsky, Salk und Glaessner (1991) bei Schülern Unterschiede zwischen den epistemologischen Überzeugungen in Bezug auf die Mathematik und die Sozialwissenschaften. Auch innerhalb einer Domäne können die epistemologischen Überzeugungen variieren (Louca et al., 2004).

Weitere Hinweise darauf, dass epistemologische Überzeugungen kontextabhängig sein können, kommen aus der Forschung zu epistemologischen Überzeugungen von Lehrern (Tobin & McRobbie, 1997, Limón, 2006b). Danach differieren die von ihnen erklärten und gezeigten¹⁶ epistemologischen Überzeugungen. Erklärte epistemologische Überzeugungen werden aus Fragebogen oder Interviews erschlossen und gezeigte epistemologische Überzeugungen werden aus dem Verhalten der Lehrer im Unterricht erschlossen. Die erklärten und gezeigten epistemologischen Überzeugungen von Lehrern sind meistens nicht identisch (Bell & Linn, 2002, Leach, Miller, Ryder & Séré, 2000, diSessa, Elby & Hammer 2002). Nach Louca et al. (2004) sind weder die erklärten noch die gezeigten epistemologischen Überzeugungen mit den zugrunde liegenden Überzeugungen einer Person identisch¹⁷. Die Autoren nehmen dabei an, dass ein Individuum weder einen direkten Zugriff zu dem deklarativen Wissen über seine epistemologischen Überzeugungen besitzt (Ansatz der Persönlichen Theorien) noch dass diese über Indikatoren (das gezeigte Verhalten, Traitansatz) erschlossen werden können. Die Kritik an der Kontextabhängigkeit in Bezug auf Traits schließt sich der in der Persönlichkeitspsychologie diskutierten Kritik von Mischel (1968) an. Danach gibt es keine transsituative Konsistenz, d. h. das Verhalten einer Person variiert in verschiedenen Situationen. Zudem lässt sich nach Mischel auch nicht das Verhalten – hier also der Umgang mit Lernmaterial – mit eigenschaftstheoretisch erfassten Maßen valide vorhersagen.

Um die genannten Probleme mit der angenommenen Konstanz epistemologischer Überzeugungen zu umgehen, schlagen Elby und Hammer vor, die Unitaritätsannahme zu verwerfen

¹⁶ In der englischen Literatur findet sich für *erklärt* der Begriff *professed* und für *gezeigt* der Begriff *enacted*. Gerade der Begriff *enacted* gibt besser als die deutsche Übersetzung wieder, dass hier die epistemologischen Überzeugungen durch Handlungen (engl. *acts*) erschlossen werden (vgl. Limón, 2006b).

¹⁷ Problematisch an der Überlegung von Louca, Elby, Hammer und Kagey (2004) ist, dass es sich um eine Annahme handelt, da die wirklichen epistemologischen Überzeugungen einer Person nicht festgestellt werden können.

und stattdessen epistemologische Überzeugungen als eine Zusammensetzung aus feiner strukturierten epistemologischen Ressourcen zu betrachten. Diese epistemologischen Ressourcen sind in Anlehnung an die *phenomenological primitives* (p-primes) von diSessa (1993) konzipiert. P-primes sind kleine Einheiten der kognitiven Architektur eines Individuums, welche das Verstehen physikalischer Vorgänge ermöglichen. Z. B. gibt es einen p-prime, um das Bewegungsverhalten eines geworfenen Balls oder die Wirkung der Kräfte beim Schieben eines Tisches zu verstehen. Je nach physikalischem Phänomen erfolgt die Aktivierung der passenden p-primes, wobei es aber auch vorkommen kann, dass ein dem Phänomen nicht entsprechender p-prime aktiviert werden kann. Im letzteren Fall liegt eine Fehlaktivierung vor, die zu einem falschen Verständnis des physikalischen Phänomens führt, wobei wichtig ist festzuhalten, dass nicht der p-prime an sich falsch ist, sondern lediglich seine Aktivierung in dem gegebenen Kontext.

Elby und Hammer postulieren nun für das Verständnis von Wissen und Wissenserwerbsprozessen wichtige kognitive Einheiten, die epistemologische Ressourcen genannt werden. Diese unterscheiden sich ontologisch von der Vorstellung epistemologischer Überzeugungen als Persönliche Theorien oder Persönlichkeitseigenschaften und sind analog zu den p-primes strukturiert. Genau wie diese sind epistemologische Ressourcen ab einer sehr frühen Entwicklungsphase aktivierbar. Ihre Aktivierung erfolgt jeweils in einem gegebenen Kontext. Damit weist das Ressourcenkonzept auch Ähnlichkeiten mit der intuitiven Physik von Sodian (2008) auf. In der intuitiven Physik werden beim Beobachten physikalischer Vorgänge von Kleinkindern kognitive Ressourcen für deren Verständnis aktiviert. Wird das Ereignis experimentell verändert (z. B. so, dass es physikalisch unmöglich wird), löst dies bestimmte Reaktionen, wie z. B. längere Blickzeiten zu dem veränderten Ereignis, aus.

Epistemologische Ressourcen bilden diejenigen Elemente der kognitiven Architektur, aus denen epistemologische Theorien oder Persönlichkeitseigenschaften zusammengesetzt sind. Die Annahme derart feingliedriger epistemologischer Ressourcen kann durch die kontextabhängige Aktivierung epistemologischer Ressourcen die Variabilität der daraus zusammengesetzten epistemologischen Überzeugungen erklären.

Elby und Hammer (2002) nehmen an, dass es vier Kategorien epistemologischer Ressourcen gibt:

- Ressourcen zum Verstehen der Natur und der Quellen des Wissens

- Ressourcen zum Verstehen epistemologischer Aktivitäten
- Ressourcen zum Verstehen epistemologischer Formen
- Ressourcen zum Verstehen epistemologischer Haltungen

In jeder dieser vier Kategorien findet sich eine Vielzahl einzelner epistemologischer Ressourcen. Die von Elby und Hammer (2002) aufgeführten werden in nachfolgender Tabelle vorgestellt und anschließend erläutert.

Tabelle 18: Beispielhafte Aufzählung epistemologischer Ressourcen nach Elby und Hammer (2002)

Ressourcen zum Verstehen der Natur und der Quellen des Wissens:

Wissen als weitergegebenes Material, Wissen als freie Kreation, Wissen als erfundener Stoff, Wissen als direkte Wahrnehmung

Ressourcen zum Verstehen epistemologischer Aktivitäten:

Akkumulation, Formation, Kontrolle, Anwendung, Vergleich, Sortierung, Benennen, Zählung, Hinzufügen

Ressourcen zum Verstehen epistemologischer Formen:

Geschichten, Regeln, Regelsysteme, Fakten, Lieder, Listen, Bilder, Kategorien, Aussagen, Wörter, Namen, Zahlen

Ressourcen zum Verstehen epistemologischer Haltungen:

Glauben, Nicht-Glauben, Zweifel, Verstehen, Verwirrung

Ressourcen zum Verstehen der Natur und der Quellen des Wissens sind kognitive Einheiten, die bestimmen, welcher Art vorliegendes Wissen ist. Z. B. kann Wissen durch direkte Wahrnehmung erworben werden oder von einer Person an eine andere Person weitergegeben werden. Die Ressourcen zum Verstehen epistemologischer Aktivitäten stellen kognitive Aktivitäten dar, mit deren Hilfe Wissen bzw. Aussagen über bestimmte Sachverhalte erschlossen werden können. Ressourcen zum Verstehen epistemologischer Formen bilden den kognitiven Grundstock, um bestimmte Formen der Wissensvermittlung richtig erkennen und verwerten zu können. So kann Wissen beispielsweise durch Geschichten oder Bilder weitervermittelt werden. Die Ressourcen zum Verstehen der epistemologischen Formen stellen entsprechend der Form der Wissensvermittlung einen dazu passenden kognitiven Verarbeitungsmechanismus bereit. Dies geht immer Hand in Hand mit Ressourcen zum Verstehen epistemologischer Aktivitäten. Z. B. erfordert das Verstehen einer erfundenen Geschichte (epistemologische Form) das Verstehen der Erschaffung des Wissens als Formation (epistemologische Aktivität), wobei zusätzlich eine Ressource zum Verstehen der Geschichte als freie Kreation (Natur und Quelle des Wissens) erforderlich ist. Die Ressourcen zum Verstehen epistemologischer Haltungen ermöglichen es, bestimmte epistemologische Haltungen einzunehmen, wie z. B. das Glauben oder Bezweifeln von Aussagen.

Die aufgezählten Ressourcen arbeiten nicht unabhängig voneinander. Das Modell epistemologischer Ressourcen ist an die kognitive Architektur von Minsky (1986) angelehnt. Nach Minsky ist das kognitive System eines Menschen aus einer Vielzahl von Agenten aufgebaut, durch deren Zusammenwirken sich die kognitiven Leistungen ergeben. Elby und Hammer lehnen sich mit ihrer Idee vom Zusammenwirken epistemologischer Ressourcen an dieses Modell an. So bewirkt die Aktivierung einer Ressource in der Regel auch die Aktivierung anderer epistemologischer Ressourcen. Wie zuvor bereits dargestellt, bedingt die Aktivierung von Ressourcen zu epistemologischen Formen notwendig die Aktivierung von Ressourcen zum Verstehen von epistemologischen Aktivitäten.

Nach Louca et al. (2004) können mit dem Ansatz epistemologischer Ressourcen die Befunde erklären werden, die sich aus entwicklungspsychologischer Perspektive und der Unitaritätsannahme ergeben. Die in entwicklungspsychologischer Perspektive auftretenden Entwicklungsstufen lassen sich durch eine regelmäßige Aktivierung der gleichen Ressourcen erklären. Louca et al. (2004) verdeutlichen dies am *Reflective-Judgment*-Modell von King und Kitcheener (1994). Hier werden die Antworten auf wissenschaftliche Fragestellungen diskutiert und aus einem Antwortschema die epistemologische Entwicklungsstufe einer Person erschlossen. Im Rahmen dieser Diskussion werden epistemologische Ressourcen aktiviert. Eine bestimmte Stufe reflexiven Denkens entspricht demnach einem bestimmten Aktivierungsmuster epistemologischer Ressourcen. Den allgemeinen Entwicklungstrend zu fortgeschrittenen Entwicklungsstufen epistemologischer Überzeugungen bzw. den Befund, dass besser gebildete Personen fortgeschrittene Entwicklungsstufen zeigen, erklären Louca et al. (2004) damit, dass durch die größer werdende Erfahrung im Umgang mit Lernen und Wissen ein immer größerer Fundus an epistemologischen Ressourcen genutzt werden kann, die zudem immer besser vernetzt werden. Zusätzlich wird die Assoziation eines bestimmten Kontexts zusammen mit der Aktivierung bestimmter epistemologischer Ressourcen immer stärker, sodass die Aktivierung dieser Ressourcen wahrscheinlicher wird.

Allerdings gibt es auch Befunde, die nahelegen, dass diese epistemologischen Entwicklungsstufen in jedem Alter zu finden sind und bei einer Person wiederkehren (Bromme, 2005). Dies wäre bei einer strikten Progression jedoch ausgeschlossen. Die Annahme von Entwicklungsstufen, die einer bestimmten Progression folgen, ist nicht unumstritten und selbst Piaget hat die „Stufen“ seines Entwicklungsmodells als Klassen des Denkens betrachtet, die keiner Sequenz unterliegen (Chapman, 1988). Diese Klassen des Denkens tauchen aber in der Ontoge-

nese eines Menschen immer in einer gleichen Sukzession auf, wobei die vorhergehenden Klassen keine zwingende Voraussetzung für die Entwicklung darauf folgender Klassen des Denkens darstellen. Die kontextabhängige Aktivierung der gleichen epistemologischen Ressourcen würde somit als eine Klasse des Denkens aufgefasst und als epistemologische Entwicklungsstufe interpretiert werden. Bedenkt man weiterhin, dass Personen sich wiederholt in ähnlichen Kontexten befinden können, kann das wiederholte Auftreten gleicher Entwicklungsstufen erklärt werden (Bromme, 2005). Somit kann dieser Ansatz auch bisher widersprüchliche Befunde erklären.

Abschließend sollen die Implikationen, welche sich aus dem Ansatz epistemologischer Ressourcen für die Forschung ergeben, kurz diskutiert werden. Die erste Folgerung betrifft die Identifikation neuer epistemologischer Ressourcen. Elby und Hammer (2002) betrachten die von ihnen aufgezählten epistemologischen Ressourcen (s. Aufzählung in Tabelle 18) nur als mögliche Beispiele. Für eine Identifikation neuer epistemologischer Ressourcen werden folgende Anforderungen gestellt: Sie müssen

- bei jungen Kindern früh in der Entwicklung festzustellen sein,
- nachvollziehbare Entwicklungsursprünge haben und
- erkennbar sein, wenn diese als „vernünftige“ Mini-Generalisierung über Wissen artikuliert werden.¹⁸

Diese Anforderungen sind als notwendige Bedingungen zur Identifikation eventuell vorzufindender epistemologischer Ressourcen zu verstehen. Wichtiger als die Ergänzung der Liste der potenziellen epistemologischen Ressourcen sind die methodischen Implikationen, die sich aus diesem Ansatz ergeben.

Die erste Folgerung betrifft die kontextabhängige Aktivierung der Ressourcen. Diese legt einen Ansatz nahe, bei dem die epistemologischen Ressourcen nicht unabhängig von dem Kontext des Lernens (oder allgemein des Wissenserwerbs) erfasst werden sollten. Allgemeine Ansätze, wie sie unter der Unitaritätsannahme zu finden sind, scheiden somit aus. Insbesondere erscheinen Fragebogenverfahren ungeeignet, um epistemologische Ressourcen zu erfassen. Geeignet für die kontextabhängige Erfassung sind insbesondere Interviews sowie die Reflexion von Alltagshandlungen. Allerdings setzt die Reflexion von Alltagshandlungen ein gewisses Entwicklungsniveau voraus, sodass diese Methode hauptsächlich für Erwachsene geeignet

¹⁸ Leider wird von Elby und Hammer nicht genauer erklärt, was unter Entwicklungsursprüngen (*developmental origins*), vernünftigen (*common-sense*) und Minigeneralisierung (*mini-generalizations*) zu verstehen ist. Insgesamt leidet der Ansatz von Elby und Hammer (2002) an seiner mangelnden sprachlichen Präzision.

erscheint. Genaue Methoden, mit denen die Reflexion von Alltagshandlungen erfasst werden soll, nennen Elby und Hammer leider nicht, verweisen aber auch hier auf Interviewverfahren. Wichtiger als die genaue Methode erscheint jedoch, dass hier versucht wird, epistemologische Ressourcen in die Nähe von tatsächlich erfolgten Handlungen zu rücken, d. h. der Fokus der Betrachtung wird von den erklärten auf die gezeigten epistemologischen Überzeugungen gerichtet.¹⁹ Weiterhin werden Fallstudien in einem naturalistischen Kontext, d. h. direkt in der Lernumgebung, und direkte Beobachtungen im Unterricht vorgeschlagen. Allen genannten Methoden ist die Kontextgebundenheit ihrer Anwendung, ausgenommen hiervon sind ist jedoch das Interviewverfahren.

Der Ansatz von Hammer und Elby stellt eine Alternative zu den bisher anzutreffenden Stufenmodellen und den persönlichkeitspsychologischen Ansätzen dar. Insbesondere wird die kognitive Grundlage epistemologischer Überzeugungen auf ein solides Fundament gestellt. Allerdings besitzt dieser Ansatz auch einige Schwächen. So ist er bisher lediglich eine phänomenale Beschreibung epistemologischer Überzeugungen, die beim Lernen hauptsächlich physikalischer Sachverhalte auftreten. Die von Hammer und Elby angeführten epistemologischen Ressourcen gliedern sich in diese phänomenale Beschreibung ein, eine funktionale Zusammenfassung der einzelnen Ressourcen in übergeordnete Einheiten erfolgt nicht. Zudem ist die Beschreibung der epistemologischen Ressourcen unspezifisch und eher oberflächlich, so dass deren Form letztendlich nicht klar wird. Auch die Kriterien, welche Anforderungen erfüllt sein müssen, um von epistemologischen Ressourcen zu sprechen, sind unklar formuliert. Nicht zuletzt ist keine Forschungsmethodik mit dem Ansatz verbunden, mit der eigenständige Erkenntnisse gewonnen werden können. Somit stellt dieser Ansatz einen ersten und auch wertvollen Schritt in Richtung einer kognitiven Architektur epistemologischer Überzeugungen dar, muss aber noch wesentlich weiter ausgebaut werden. Ein anderes, älteres und nicht sehr bekanntes Modell epistemologischer Überzeugungen, das ebenfalls den kognitiven Aufbau behandelt, stammt von Royce (1978) und wird im Folgenden neben anderen alternativen Ansätzen vorgestellt.

1.5 Alternative Ansätze epistemologischer Überzeugungen

Im Folgenden werden noch einige Sichtweisen epistemologischer Überzeugungen vorgestellt, die nach der Klassifikation aus Tabelle 1 den alternativen Ansätzen zuzurechnen sind. Bei

¹⁹ Unter epistemologischer Überzeugung ist hier das Ensemble der im jeweiligen Kontext aktivierten epistemologischen Ressourcen zu verstehen.

diesen Modellen stehen nicht die Eigenschaften epistemologischer Überzeugungen im Vordergrund, sondern deren Einbettung in den instruktionalen Kontext sowie in die kognitiven Funktionen. Von diesen Sichtweisen wird zunächst die *Psychologische Epistemologie* von Royce (1978) vorgestellt. Royce erarbeitet, ausgehend von den grundlegenden kognitiven Prozessen, ein Modell über interdisziplinäre Unterschiede epistemologischer Überzeugungen. Das TIDE-Modell von Muis et al. (2006) beschreibt die allgemeine Entwicklung epistemologischer Überzeugungen und die vertikale Differenzierung, ausgehend von allgemeinen epistemologischen Überzeugungen bis zu speziellen, kontextbezogenen epistemologischen Überzeugungen. Zum Abschluss wird das Modell von Hofer (2004b) vorgestellt, welches epistemologische Überzeugungen als Metawissen (Wissen über das Wissen) beschreibt und die Dimensionen von Hofer und Pintrich (1997) entsprechend verortet.

1.5.1 Psychologische Epistemologie von Royce

Die psychologische Epistemologie von Royce (1978) ist ein Ansatz, der zeitgleich mit dem Stufenmodell von Perry (1970) entwickelt wurde, aber niemals dessen Bekanntheit erlangte. Im Folgenden wird die Psychologische Epistemologie in Anlehnung an die Darstellung in Muis et al. (2006) vorgestellt. Der Ansatz von Royce richtet sich in erster Linie nicht auf die Entwicklung epistemologischen Denkens oder dessen dimensionale Struktur, sondern konzentriert sich auf die kognitiven Prozesse, die zum Erwerb von Wissen notwendig sind. Nach Royce ist das Verstehen der Prozesse, auf welche Art und Weise ein Individuum Wissen erlangt, notwendig zum Verständnis epistemologischer Fragen. Von einer psychologischen Perspektive wird also Bezug genommen auf eine philosophische Perspektive. Wissen wird dabei als die Kognitionen eines Individuums definiert, die epistemologisch zu rechtfertigen sind. Der psychologisch interessierende Bereich wird somit durch ein philosophisches Kriterium eingegrenzt.

Royce (1959) unterscheidet *Modi des Wissens*. Danach ist z. B. Wissen in der Wissenschaft etwas anderes als Wissen in der Kunst. Nach der bereits aufgeführten Definition von Wissen unterscheiden sich somit die kognitiven Prozesse, die zur epistemologischen Rechtfertigung von Aussagen in der Wissenschaft und in der Kunst benutzt werden. In epistemologischer Hinsicht liegen unterschiedliche Kriterien der Rechtfertigung vor und aus psychologischer Sicht werden unterschiedliche kognitive Prozesse im jeweiligen Kontext in unterschiedlichem Ausmaß angewandt. Royce bezeichnet dies als *Way of Knowing*. In verschiedenen Bereichen bildet sich im Verlauf der bereichsspezifischen Sozialisation eine spezielle Präferenzordnung

dieser *Ways of Knowing* heraus, es ergibt sich eine hierarchische Struktur der Prozesse, die zur Rechtfertigung von Aussagen benutzt werden (Royce, 1978).

Nach Royce (1978) sowie Royce und Mos (1980) lassen sich grundsätzlich drei *Ways of Knowing* unterscheiden. Dies sind Rationalismus, Empirismus und Metaphorismus. Rationalismus bedeutet, dass Wissen durch logisches, konzeptuelles und analytisches Denken erworben wird. Grundlegend hierfür ist die logische Konsistenz. Beim Empirismus wird Wissen durch wiederholte und strukturierte Beobachtungen und Daten gewonnen. Beim Metaphorismus rückt die Subjektivität von Wissen in den Mittelpunkt. Hier wird konstatiert, dass Wissen persönlich ist und der Integration von Aussagen und des Gebrauchs von Symbolen bedarf. Diese *Ways of Knowing* sind nun direkt abhängig von den kognitiven Prozessen, die zur Rechtfertigung von Wissen gebraucht werden. Im Laufe der Sozialisation im Bildungssystem und dem Fortschreiten in einer bestimmten Domäne bilden sich spezielle Formen epistemologischen Denkens heraus (Royce, 1978). Die Ausbildung stellt das Fortschreiten einer Person vom Status eines Novizen zum Status eines Experten in einer bestimmten Domäne dar. Während der Novize nicht über spezialisiertes Wissen verfügt und dieses Wissen nicht abhängig von der jeweiligen Domäne ist, besitzt der Experte spezialisierte Wissensformen, die abhängig von der vorherrschenden Epistemologie einer Domäne sind. Je nach Domäne werden dabei bestimmte epistemologische Denkweisen bevorzugt und es ergibt sich als Folge der Ausbildung die oben genannte Präferenzordnung der *Ways of knowing*.

Diese Präferenzordnung wird als psychoepistemologisches Profil bezeichnet. Zur Bestimmung des psychoepistemologischen Profils einer Person wurde der PEP-Fragebogen entwickelt (Royce & Smith, 1964, Royce & Mos, 1980). Der Fragebogen besteht aus drei Skalen, wobei jede Skala einen bestimmten *Way of Knowing* misst. Die Skala mit dem höchsten Wert gibt die bevorzugte epistemologische Denkweise einer Person wieder. Die Skalen des PEP grenzen sich gegenüber anderen kognitiven Variablen ab und es hat sich gezeigt, dass die Skalen des PEP bei einer gemeinsamen Faktorenanalyse mit anderen Maßen kognitiver Fähigkeiten einen eigenständigen Faktor bilden (Mos, Wardell & Royce, 1974).

Um zu überprüfen, ob sich die dominanten epistemologischen Denkweisen wie angenommen zwischen Domänen unterscheiden, haben Royce und Mos (1980) den PEP Universitätsprofessoren aus den Bereichen Literatur und Kunst, Mathematik und theoretische Physik sowie Naturwissenschaften vorgelegt. Bei den Professoren aus dem literarisch-künstlerischen Bereich

war der Metaphorismus die vorherrschende epistemologische Denkweise, im Bereich der Mathematik und theoretischen Physik der Rationalismus und bei den Professoren der Naturwissenschaften der Empirismus.

In Weiterführung dieses Ansatzes hat Donald (1990) Unterschiede zwischen Experten in verschiedenen Domänen untersucht. Dabei wurde zur Unterscheidung von Domänen das Klassifikationsschema von Biglan (1973a, b) verwendet. Danach lassen sich akademische Domänen nach dem Vorhandensein eines gemeinsamen Paradigmas, im Hinblick auf die praktische Anwendung und der Beschäftigung mit lebenden Systemen aufgliedern, wobei die Beschäftigung mit lebenden Systemen ein sehr schwaches Unterscheidungskriterium ist. Damit ergeben die in Tabelle 19 zusammengestellten Gliederungsmöglichkeiten.

Tabelle 19: Klassifikation akademischer Domänen nach Biglan (1973b)

	Akademische Domäne	
Existenz eines Paradigmas	hart	weich
praktische Anwendung	rein	angewandt
Beobachtung lebender Systeme	ja	nein

Diese Unterscheidungskriterien stellen keine disjunkten Kategorien dar, es gibt immer jeweils einen Überschneidungsbereich. Beispielsweise wird in der Mathematik zwischen reiner und angewandter Mathematik unterschieden. Allerdings ist die Mathematik wegen ihres paradigmatischen Aufbaus sowohl in reiner als auch angewandter Perspektive als hart zu bezeichnen, etwa im Gegensatz zur Psychologie, die kein einheitliches Paradigma besitzt, sondern in der eher eine Vielfalt an Paradigmen vorherrscht. Eine weitere Gliederungsmöglichkeit stellt die Dimension *gut strukturiert* vs. *schlecht strukturiert* von Frederikson (1984) dar. Diese Dimension bezieht sich darauf, ob eine bestimmte Domäne identifizierbare Lösungsstrategien besitzt (wie z. B. die gut strukturierte Mathematik) oder sich auf idiosynkratische und fallweise Lösungsansätze beruft (wie z. B. die schlecht strukturierten Sozialwissenschaften). Donald (1990) verglich nun Professoren der Natur-, Sozial- und Geisteswissenschaften in Bezug auf die vorherrschende Epistemologie. Bei Natur- und Sozialwissenschaftlern fand sich die Orientierung an der empirischen Evidenz als die vorherrschende epistemologische Denkweise, während sich Geisteswissenschaftler zur Rechtfertigung ihrer Aussagen auf externe Autoritäten verlassen (z. B. im Zuge eines Peer Review). Donald fand überdies, dass sich Experten in einer Domäne nicht nur auf eine bestimmte epistemologische Denkweise verlassen. Insbesondere gibt es auch innerhalb einer Domäne Unterschiede in der vorherrschenden epistemologischen Denkweise, wenn sich der Bezug von der reinen Wissenschaft auf eine angewandte

Perspektive ändert. So verwenden Sozialwissenschaftler, die sich mit rein wissenschaftlichen Fragen beschäftigen, empirische Evidenz sowie Gegenbeispiele und alternative Erklärungen (Rationalismus) als vorherrschende epistemologische Denkweise, während Sozialwissenschaftler, die sich mit angewandten Fragestellungen beschäftigen, sich nur auf empirische Evidenz verlassen. Offensichtlich gibt es auch innerhalb einer akademischen Domäne eine gewisse Variabilität, was die Verwendung epistemologischer Denkweisen anbelangt. Gliedert man die Domänen nach der Dimension gut strukturiert (Mathematik) – schlecht strukturiert (Psychologie), so ergibt sich zudem, dass es in den gut strukturierten Domänen weniger Variabilität in der Verwendung epistemologischer Denkweisen gibt als in schlecht strukturierten Domänen.

Vergleiche sind zwischen allen möglichen Disziplinen in allen der vier oben genannten Klassifikationsmöglichkeiten denkbar. Royce (1978) zieht Vergleiche im Hinblick auf die Klassifikationsaspekte Struktur, Existenz eines gemeinsamen Paradigmas und praktische Anwendung und die vorherrschende epistemologische Denkweise.

Tabelle 20: Vergleich verschiedener akademischer Domänen nach Royce (1978)

Disziplin	Strukturierung	Existenz eines Paradigmas	praktische Anwendung	vorherrschende epistemologische Denkweise
Physik	gut strukturiert	hart	rein	Empirismus
Psychologie	schwach strukturiert	weich	angewandt	Empirismus
Erziehungswissenschaft	schwach strukturiert	weich	angewandt	Empirismus
Geschichte	schwach strukturiert	weich	rein	Metaphorismus
Mathematik	gut strukturiert	hart	rein	Rationalismus
Ingenieurwissenschaften	gut strukturiert	hart	angewandt	Empirismus

Zusammenfassend ist festzuhalten, dass sich beim Vergleich zwischen Domänen mittels des Ansatzes von Royce Unterschiede in den vorherrschenden epistemologischen Denkweisen auffinden lassen. Das Besondere am Ansatz von Royce ist, dass direkt auf kognitive Prozesse Bezug genommen wird, die zur Rechtfertigung von Aussagen benutzt werden und nur solche gerechtfertigte Aussagen als Wissen gelten. Der Ansatz von Royce sowie die darauf basierenden Arbeiten machen deutlich, dass Personen über eine kognitive Ausstattung zur Rechtfertigung von Aussagen verfügen. Im Verlauf der Sozialisation in einer bestimmten Domäne bil-

det sich dann eine Präferenzordnung dieser Prozesse aus, sodass bestimmte Rechtfertigungsprozesse gegenüber anderen bevorzugt werden.

1.5.2 Theory of Integrated Domains (TIDE) – Framework

Das TIDE-Framework von Muis et al. (2006) beschreibt die Einbettung epistemologischer Überzeugungen in den allgemeinen, soziokulturellen Kontext einer Person sowie die individuelle Ontogenese während der ganzen Lebensspanne eines Menschen. Im TIDE-Framework werden epistemologische Überzeugungen als situiert im soziokulturellen Kontext eines Individuums betrachtet. Die Entwicklung allgemeiner epistemologischer Überzeugungen beginnt bei der Geburt und endet mit dem Tod eines Individuums. Epistemologische Überzeugungen sind in mehrere vertikale Schichten untergliedert, zwischen denen es Wechselwirkungen gibt. Auf der untersten Ebene finden sich die allgemeinen epistemologischen Überzeugungen, auf der darüber liegenden Ebene akademische epistemologische Überzeugungen und auf der dritten und spezifischsten Ebene diejenigen epistemologischen Überzeugungen, die sich auf eine bestimmte Domäne und den dazugehörigen instruktionalen Kontext beziehen.

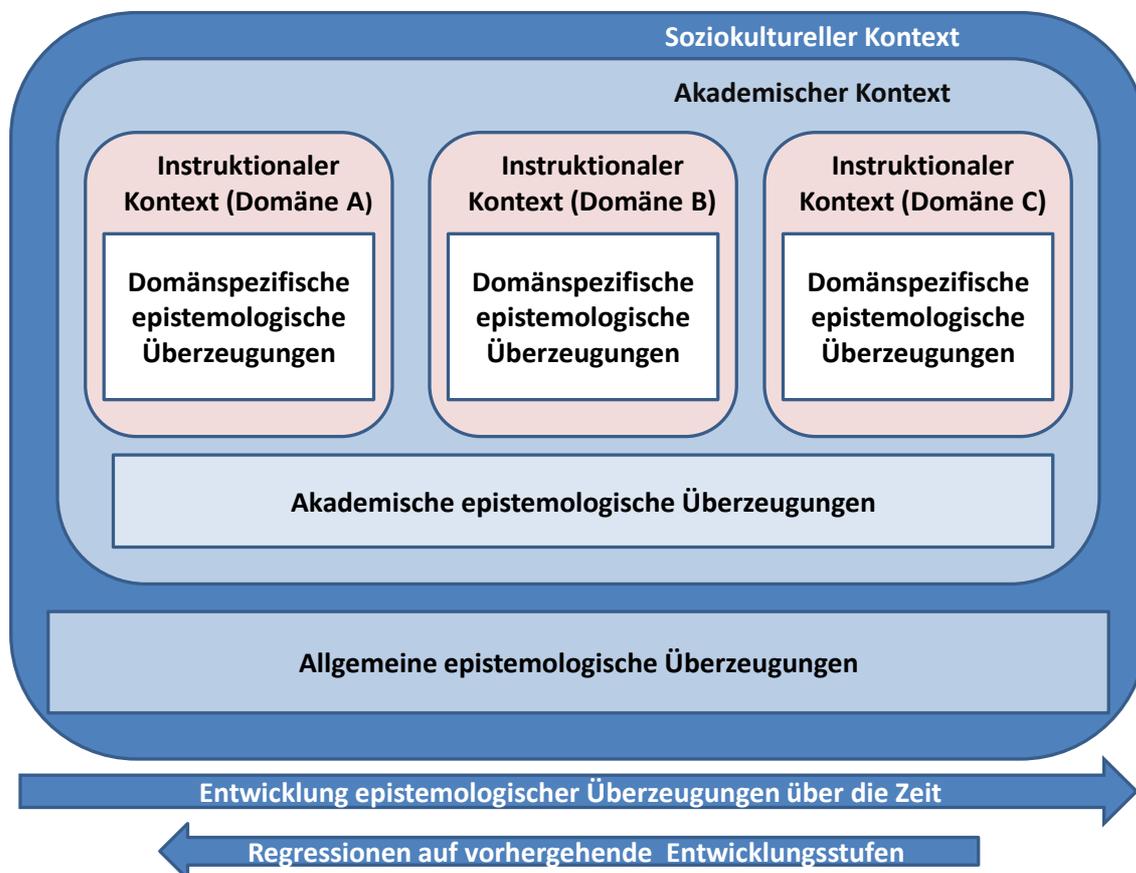


Abbildung 5: TIDE-Framework von Muis et al. (2006), dargestellt für die drei Domänen A, B und C

Beim TIDE-Framework werden die Individuen in ihrem jeweiligen soziokulturellen Kontext betrachtet. In diesem Kontext vollzieht sich die Entwicklung der allgemeinen epistemologischen Überzeugungen. Entwicklung wird als Interaktion des Individuums mit der Umwelt verstanden (Fischer, 1980) und ist somit auch konform mit den Vorstellungen Piagets über die individuelle Entwicklung (Chapman, 1988). Die Aneignung allgemeiner epistemologischer Überzeugungen wird als Prozess der Enkulturation verstanden (Gay, 1978, Pai, 1990, Jehng et al., 1993, Paulsen & Wells, 1998, Hofer, 2000, Estes, Chandler, Horvath & Backus, 2003, Schommer-Aikins, Duell & Barker, 2003). In diesem soziokulturellen Kontext vollzieht sich die Entwicklung allgemeiner epistemologischer Überzeugungen. Diese entwickeln sich in nicht akademischen Kontexten in der jeweiligen Interaktion mit Peers, verschiedenen Arbeitsumgebungen oder ähnlichen Umweltbedingungen. Dabei liegt ein evolutionäres Konzept zugrunde, d. h. Kinder verfügen zu Beginn der Entwicklung über intuitive Theorien (vgl. Sodian, 2008), denen ein angeborener kognitiver Mechanismus zugrunde liegt. Ausgehend von diesen intuitiven Theorien entwickeln Kinder naive Konzepte des Wissens, aus denen später allgemeine epistemologische Überzeugungen werden. Die auf diesen naiven Wissenskonzepten beruhende Wissensaneignung ist informell und findet im Alltagskontext eines Individuums statt. Sie ist von der formellen Wissensaneignung in einem akademischen Kontext zu unterscheiden. Die formelle und die informelle Wissensaneignung können sich gegenseitig behindern, müssen es jedoch nicht. Dass es beide Arten der Wissensaneignung gibt, wurde von Schoenfeld (1985) bestätigt. Es zeigte sich, dass Schüler formale Aufgaben aus der Schulmathematik nicht lösen konnten, während sie in einen Alltagskontext eingebettete gleichartige Aufgaben lösen konnten. Hier wird insbesondere deutlich, dass dieses Konzept an evolutionspsychologische Überlegungen anschließt. Cosmides (1989) sowie Cosmides und Tooby (1992) konnten nachweisen, dass Versuchspersonen in dem auf logischen Schlussfolgerungen beruhenden *Wason Selection Task* (Wason, 1966, 1968) deutlich bessere Ergebnisse zeigen, wenn die Aufgaben im Kontext des sozialen Austauschs dargeboten wurden. Cosmides und Tooby (1992) interpretieren dieses Ergebnis evolutionspsychologisch. Menschen besitzen demzufolge evolvierte psychologische Mechanismen,²⁰ welche darauf ausgelegt sind,

²⁰ Ein evolviertes psychologischer Mechanismus (EPM) ist ein Prozess oder eine Menge von Prozessen, die sich durch folgende Eigenschaften charakterisieren lassen (Schwab, 2004): 1) Er existiert in der gegebenen Form, weil er zur Lösung eines speziellen Anpassungsproblems beiträgt. 2) Er reagiert und verarbeitet nur eine ganz bestimmte Art von Information. 3) Der Input eines EPM gibt dem Organismus eine Information über das spezielle Anpassungsproblem, dem der Organismus begegnen muss. 4) Der Input eines EPM wird durch Entscheidungsregeln (Heuristiken) in einen Output transformiert. 5) Der Output eines EPM kann aus physiologischer Aktivität, manifestem Verhalten oder wiederum aus Input für einen anderen EPM bestehen. 6) Der Output eines EPM ist auf die Lösung des spezifischen Anpassungsproblems gerichtet, das zu der Entstehung des EPM führte. Ein EPM lässt sich im Sinn der kognitiven Psychologie als Blackbox bezeichnen. Aus den o. g. Eigenschaften eines EPM lassen sich noch einige zusätzliche Eigenschaften ableiten. So sind EPM bereichsspezifisch, d. h. für

speziell diejenigen Aufgaben zu bearbeiten, die im Kontext des sozialen Austauschs entstehen. Die Übertragung auf kontextferne Aufgaben ist nicht möglich. Diese Ergebnisse legen nahe, dass die Informationsverarbeitung nicht unabhängig vom Inhalt geschieht. Ähnlich wie diese evolvierten psychologischen Mechanismen für den sozialen Austausch sind auch die Mechanismen, welche eine grundlegende Verarbeitung von Wissen gestatten. Eine notwendige Folge daraus ist, dass diese Mechanismen nicht unabhängig von ihrem Inhalt funktionieren und eine Übertragung auf andere Inhalte nicht möglich ist.

Eingebettet in den soziokulturellen Kontext einer Person ist der akademische Kontext, unter dem jede Form der formellen Wissensvermittlung wie schulische und universitäre Wissensvermittlung verstanden wird. Im akademischen Kontext entwickeln sich die akademischen epistemologischen Überzeugungen und später, eingebettet in einen speziellen instruktionalen Kontext, die domänenspezifischen epistemologischen Überzeugungen. Die Entwicklung akademischer epistemologischer Überzeugungen beginnt beim Eintritt in das Bildungssystem. Zu Beginn sind die akademischen epistemologischen Überzeugungen nicht oder nur wenig von den allgemeinen epistemologischen Überzeugungen verschieden. Ausgehend von diesen allgemeinen epistemologischen Überzeugungen beginnt die Entwicklung akademischer epistemologischer Überzeugungen, die sich auf fachübergreifende Überzeugungen zur Struktur des Wissens und zu den Prozessen des Wissenserwerbs beziehen. Im Verlauf der akademischen Enkulturation erfolgt die Spezialisierung dieser akademischen epistemologischen Überzeugungen zu domänenspezifischen epistemologischen Überzeugungen. Domänenspezifische epistemologische Überzeugungen wirken aber ihrerseits auf die akademischen Überzeugungen zurück, sodass ein komplexes Wirkungsgefüge entsteht. Im Entwicklungsverlauf entstehen auf diese Weise epistemologische Überzeugungen, die Agglomerate der allgemeinen epistemologischen Überzeugungen, den früher ausgebildeten akademischen Überzeugungen sowie den jeweiligen domänenspezifischen epistemologischen Überzeugungen darstellen.

Die Entwicklung der domänenspezifischen epistemologischen Überzeugungen vollzieht sich in einem jeweils besonderen instruktionalen Kontext. Im Zuge der Spezialisierung auf ein Fach wird eine bestimmte domänenspezifische epistemologische Überzeugung dominant. Empirische Unterstützung für diese Annahme liefern die bei der Vorstellung der psychoepistemologischen Profile angeführten Befunde von Donald (1990). Domänenspezifische epistemologische

eine bestimmte Art von Anpassungsproblem existiert jeweils ein bestimmter EPM, der zur Lösung des Anpassungsproblems beiträgt. Dadurch folgt zusätzlich, dass Menschen eine große Menge an EPM besitzen müssen. Die große Anzahl und Bereichsspezifität verleiht dem menschlichen Verhalten große Flexibilität.

Überzeugungen beziehen sich somit auf eine genau definierte Domäne. Dabei beeinflusst die instruktionale Umgebung die Entwicklung der jeweiligen domänenspezifischen epistemologischen Überzeugungen, d. h. im TIDE-Framework wird eine Wirkrichtung von der Umwelt auf das Individuum unterstellt. Die domänenspezifischen epistemologischen Überzeugungen (und durch die bestehenden Wechselwirkungen auch die akademischen epistemologischen Überzeugungen) sind eine subjektive Rekonstruktion einer Person, auf welche Art und Weise Wissen im jeweiligen Fachgebiet betrachtet wird. Dieses „Modell des Wissens“ und dessen Bedeutung werden durch die soziale Interaktion im instruktionalen Kontext ausgehandelt (Bauersfeld, 1988). Es besteht also eine Interaktion zwischen Lehrenden und Lernenden. Die Vorstellung ist analog zu dem Wirkmodell von Hofer (2001).

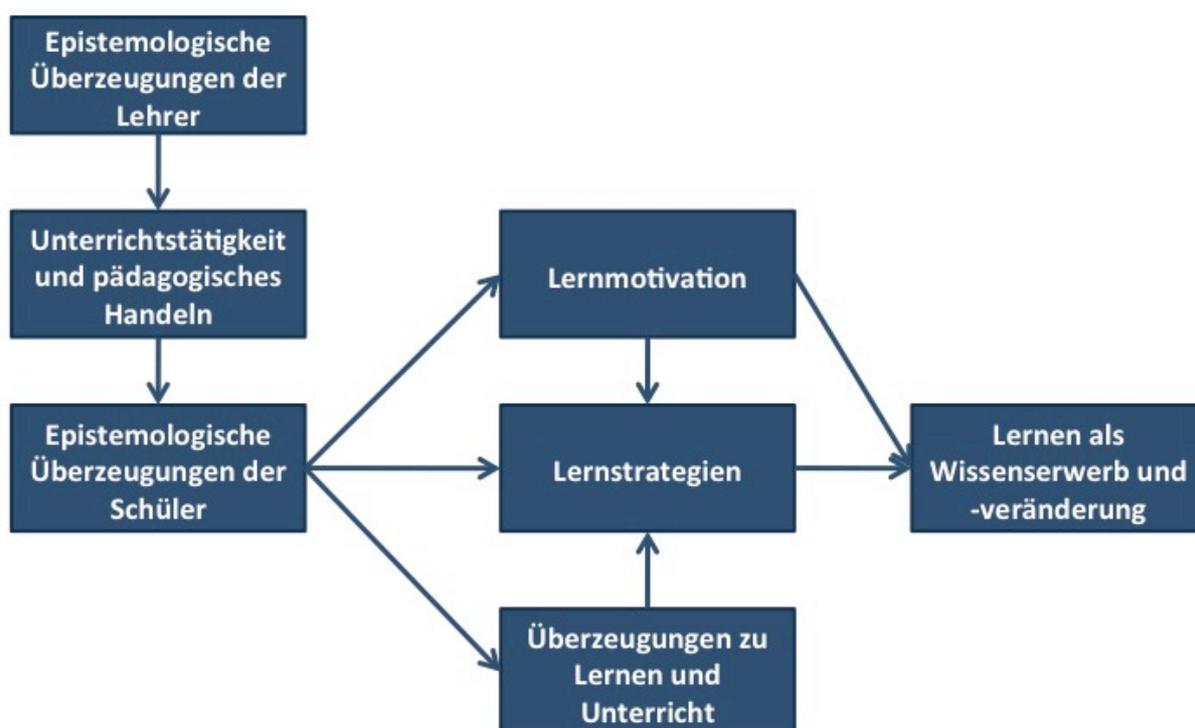


Abbildung 6: Wirkmodell von Hofer (2001)

Nach diesem Wirkmodell beeinflussen die epistemologischen Überzeugungen der Lehrkräfte deren Handeln im Unterricht, was wiederum einen Einfluss auf die epistemologischen Überzeugungen der Schüler hat.²¹ Muis et al. (2006) nehmen an, dass es sich bei den entstehenden epistemologischen Überzeugungen der Schüler um die gezeigten epistemologischen Überzeugungen der Lehrer handelt (Louca et. al, 2004)²². Eine Folge davon ist, dass die von der Per-

²¹ Hofer (2001) spricht in dem Modell von epistemologischen Theorien und macht damit Gebrauch von dem Ansatz epistemologischer Überzeugungen als Persönliche Theorien.

²² Wobei sicherlich noch andere Faktoren als die Lehrer die epistemologischen Überzeugungen von Schülern beeinflussen.

son geäußerten epistemologischen Überzeugungen nicht mit der von ihr in einem konkreten Kontext gezeigten übereinstimmen müssen.

Das TIDE-Framework bietet also eine mehrschichtige Beschreibung epistemologischer Überzeugungen und die Möglichkeit, diese in den jeweils interessierenden Kontext einzuordnen. Insbesondere werden in dem Rahmenwerk die Entwicklung epistemologischer Überzeugungen sowie deren Kontextabhängigkeit berücksichtigt. Dabei beschäftigt sich das TIDE-Framework nicht mit der internen Struktur epistemologischer Überzeugungen, sondern mit deren Entwicklung und Verortung im soziokulturellen, akademischen und instruktionalen Umfeld des Individuums.

Aus diesem Rahmenwerk ergeben sich auch methodische Schlussfolgerungen. Durch die kontextuelle Einbettung epistemologischer Überzeugungen wird die Frage aufgeworfen, ob jeweils auf den spezifischen Kontext bezogene Fragebogen zur Erfassung epistemologischer Überzeugungen oder allgemeine Fragebogen angewandt werden sollen. Muis et al. (2006) sehen die Entscheidung über den Einsatz allgemeiner oder kontextabhängiger Fragebögen abhängig von der Fragestellung. Will man beispielsweise den Zusammenhang mit Durchschnittsnoten über verschiedene Fächer hinweg untersuchen, bietet sich die Erfassung akademischer epistemologischer Überzeugungen an, da diese ähnlich wie die Durchschnittsnote ein aggregiertes Maß über verschiedene Fächer (und damit auch instruktionale Kontexte) darstellen. Wichtig ist, dass der Spezifizierungsgrad der Variablen in der aufgestellten Hypothese übereinstimmt. Ist z. B. die abhängige Variable wie im obigen Beispiel ein aggregiertes Maß, kann die unabhängige Variable auch aggregiert sein. Bezieht sich die abhängige Variable auf eine genau umrissene Variable in einer Disziplin, beispielsweise die Lernleistung im Fach Mathematik, so ist auch die Erfassung der domänenspezifischen epistemologischen Überzeugungen notwendig.

Das zuletzt dargestellte Modell epistemologischer Überzeugungen beschreibt ähnlich wie das TIDE-Framework nicht die Binnenstruktur epistemologischer Überzeugungen, sondern behandelt die Einbettung epistemologischer Überzeugungen in das kognitive System, insbesondere den Bezug epistemologischer Überzeugungen zur Metakognition. Von besonderem Interesse ist dabei nicht nur diese Verortung epistemologischer Überzeugungen im kognitiven System des Menschen, sondern auch die Methodik, mit welcher hier epistemologische Überzeugungen erfasst werden.

1.5.3 Epistemologische Überzeugungen und Metakognition

Hofer (2004b) beschreibt epistemologische Überzeugungen als Teil des metakognitiven Systems eines Menschen. Diese sind somit ein Teil der metakognitiven Prozesse eines Individuums, die mit der Planung, Steuerung und Kontrolle der kognitiven Prozesse befasst sind. So verstandene epistemologische Überzeugungen stellen Wissen über das Wissen dar. Dies ist analog zur Metakognition, die umgangssprachlich als Denken über das Denken umschrieben werden kann. Epistemologische metakognitive Funktionen stehen nach Kitchener (1983) auf der obersten hierarchischen Ebene des kognitiven Systems. Die Stufen der *kognitiven Pyramide* sind in Abbildung 7 dargestellt.



Abbildung 7: Kognitive Pyramide nach Kitchener (1983)

Auf der ersten Ebene finden sich die normalen kognitiven Funktionen wie Wahrnehmen, Verarbeiten, Speichern, Abrufen etc. Die zweite Ebene wird durch die metakognitiven Funktionen gebildet. Ihre Aufgabe ist die Überwachung der Funktionen der ersten Ebene, darunter fällt z. B. das Wissen über die Anwendung der einzelnen kognitiven Funktionen sowie das Wissen über die strategische Anwendung der einzelnen Funktionen. Auf der dritten Ebene findet sich die epistemische Kognition, welche das Wissen über die Sicherheit und Grenzen des Wissens beinhaltet sowie die Kriterien, wann Wissen als gültig erachtet werden kann. Jede übergeordnete Stufe benötigt für ihre Funktion die vorhergehende Stufe, wobei die Funktionen der unteren Stufe nicht in die Funktionen der übergeordneten Stufe eingeschlossen sind.

sen sind. Übrigens ist es nicht erforderlich, dass die metakognitiven Funktionen einer Person bewusst sind (Schraw & Moshmann, 1995). Analog zu diesem dreiteiligen Modell spricht Kuhn (1999) in ihrem Ansatz des Metawissens ebenfalls von drei Hierarchiestufen. Auf der ersten Stufe findet sich metakognitives Wissen. Hier findet die Kontrolle der kognitiven Funktionen statt, die sich auf deklaratives Wissen beziehen. Auf der zweiten Stufe findet sich metastrategisches Wissen, das die Kontrolle über das prozedurale Wissen einer Person ausübt. Auf der dritten und letzten Hierarchiestufe findet sich epistemologisches Metawissen, das Wissen über Wissen und Wissenserwerb beinhaltet.

Epistemologische Funktionen lassen sich somit als ein Teil der metakognitiven Aktivitäten einer Person auffassen. Von Interesse ist nun, auf welche Art und Weise epistemologische Überzeugungen mit ihren unterschiedlichen Funktionen in die metakognitiven Aktivitäten eingegliedert werden können. Hofer (2004b) benutzt dazu das Modell der Metakognition von Pintrich, Wolters und Baxter (2000), welches wie die beiden zuvor dargestellten Modelle drei Komponenten vorsieht: metakognitives Wissen, metakognitive Beurteilung und Überwachung sowie Selbstregulation und Kontrolle der Kognition. Metakognitives Wissen stellt das Wissen einer Person über ihre eigenen kognitiven Funktionen und Strategien dar und beinhaltet auch das Wissen um Faktoren, welche die Funktionen des kognitiven Systems beeinflussen können. Zusätzlich kommt das Wissen um das Selbst als denkende und lernende Person hinzu. Diese Komponente ist relativ statisch. Metakognitive Beurteilung und Überwachung stellen eine prozessorientierte Komponente dar und beinhalten das Einschätzen der Aufgabenschwierigkeit, die Überwachung des eigenen Verstehens und Lernens sowie das Einschätzen der Sicherheit der Problemlösung. Selbstregulation und Kontrolle der Kognition beinhalten die Planung, Ressourcenallokation, Strategiewahl und volitionale Kontrolle. In diesem Dreikomponenten-Modell von Pintrich et al. (2000) werden die vier Dimensionen epistemologischer Überzeugungen nach Hofer und Pintrich (1997) eingegliedert, diese Eingliederung ist in Tabelle 21 dargestellt.

Hofer illustriert die Integration epistemologischer und metakognitiver Prozesse am Beispiel einer Person, die widersprüchliche Informationen (sei es im Sinne einer Theorie und widersprechender Daten oder zweier konkurrierender Theorien) verarbeiten muss. Im Sinne der metakognitiven Kontrolle muss die Person bestimmen, welche der Wissensquellen sie als valide wahrnimmt und welche Rechtfertigung den jeweiligen Argumenten bzw. Daten zugrunde liegt. Diese Beurteilungsprozesse sind abhängig von der jeweiligen Erfahrung im Um-

gang mit solchen Situationen bzw. von der Entwicklungsstufe der epistemologischen Überzeugung. Im Sinne metakognitiven Wissens muss die Person die Sicherheit der jeweils konkurrierenden Theorien oder Daten beurteilen. Hierbei spielen alle Faktoren eine Rolle, welche die Informationsverarbeitung der Person beeinflussen, wie beispielsweise deren *need for cognition*. Überdies werden bei der metakognitiven Beurteilung und Überwachung die jeweiligen konkurrierenden Argumente inhaltlich beurteilt, wohingegen die metakognitive Kontrolle der Sicherheit des Wissens nicht vom Inhalt abhängt (Weinstock, 2006). Weiterhin bedient sich die Person einer Kontrolle ihrer kognitiven Aktivitäten während der Wissenskonstruktion. Diese findet abhängig von determinierenden Einflussgrößen wie motivationalen oder kognitiven Variablen, z. B. Leistungsmotivation und Denkstilen, statt.

Tabelle 21: Epistemologische Überzeugungen und metakognitive Komponenten nach Hofer (2004b)

Metakognitive Komponente	Dimensionen epistemologischer Überzeugungen
Metakognitives Wissen	<ul style="list-style-type: none"> • <i>Sicherheit des Wissen</i>: Annahmen, dass Wissen sicher oder nur zeitweilig und veränderbar ist • <i>Einfachheit des Wissens</i>: Annahmen, dass Wissen eine Sammlung von Fakten ist und ob diese Fakten komplex und verbunden sind • <i>Selbst als denkende und lernende Person</i>: Alle Persönlichkeitseigenschaften, die mit epistemologischen Überzeugungen zusammenhängen, wie z. B. <i>need for cognition</i>
Metakognitive Beurteilung und Überwachung	<ul style="list-style-type: none"> • <i>Beurteilung der Wissensquelle</i>: „Woher weiß ich das?“, „Passt dieses Wissen und meine Erfahrungen zusammen?“ • <i>Bestimmung der Rechtfertigung des Wissens</i>: „Ist dies glaubwürdig?“, „Wird die Behauptung durch Fakten gestützt?“, „Kann man Daten und Theorie abgleichen?“
Selbstregulation und Kontrolle der Kognition	<ul style="list-style-type: none"> • <i>Regulation der Kognition während der Wissenskonstruktion</i>: Selbstregulation des Wissenserwerbs wird durch Interessen, Motivation, Denkstile, Werte und Überzeugungen gesteuert, „Weiß ich, was ich wissen muss, oder brauche ich mehr Wissen?“

Ein weiterer Aspekt, der sich aus der Betrachtung epistemologischer Überzeugungen als Teil der metakognitiven Aktivität eines Individuums ergibt, sind die mit diesem Ansatz verbundenen methodischen Implikationen. Insbesondere im Zusammenhang mit der Verwendung von Fragebögen weist Hofer auf mögliche Probleme mit der Dimensionalität des Konstrukts hin. Bei epistemologischen Dimensionen, die sich nicht klar auf einem Kontinuum zwischen zwei Polen abtragen lassen, ergeben sich Schwierigkeiten bei der Benutzung von Ratingskalen.²³

²³ Hofer (2004b) spricht an dieser Stelle von der Verwendung von Likert-Skalen. Hier sei noch einmal darauf hingewiesen, dass es sich bei einer Likert-Skala immer um eine Menge von Items handelt, von denen jedes mit-

Hofer führt als Beispiel die Dimension *Wissensquelle* an. Bei dieser ist es besonders schwer, die beiden Endpunkte zu bestimmen, da sich die Pole im Verlauf der Entwicklung mehrfach ändern. Das Konstrukt *Wissensquelle* ist damit zu komplex, um durch zweidimensionale Skalen erfasst zu werden. Ein weiterer Punkt betrifft die der Fragebogenmethode zugrunde liegende Annahme, dass epistemologische Überzeugungen dem Individuum bewusst zugänglich sind und unabhängig vom jeweiligen Kontext abgefragt werden können. Die Kritik von Hofer geht damit in die gleiche Richtung wie die zuvor dargestellte Kritik von Hammer und Elby (2002).

Weiter stellt sich die Frage, welche methodischen Alternativen sich nun aus der metakognitiven Sichtweise epistemologischer Überzeugungen ableiten lassen? Hofer sieht in den Methoden der kognitiven Psychologie zur Erfassung von Metakognition eine Quelle, die in der Forschung zu epistemologischen Überzeugungen vorzufindenden Methoden anreichern zu können. Pintrich et al. (2000) führen dabei Strategiefragebogen, Multiple-Choice-Tests, Error-detection-Studien, Interviews, Selbstberichtverfahren sowie lautes Denken als mögliche Zugänge an. Interviewverfahren sind seit den Anfängen bei Perry Bestandteil des Methodeninventars. Hofer sieht das laute Denken als eine besonders geeignete Methode zur Erfassung der metakognitiven Aspekte epistemologischer Überzeugungen. Schraw (2000) gibt jedoch mögliche Verzerrungen an, die durch lautes Denken verursacht werden können. Zum einen kann die Aufmerksamkeit des Probanden zu sehr auf das laute Denken gelenkt werden und zum anderen kann der Hinweis zum lauten Denken zu *demand characteristics* führen, d. h. der Proband versucht, im Sinne der Untersuchung zu antworten. Obwohl lautes Denken auch eine Reihe von Schwachstellen erkennen lässt, bietet es eine gute Möglichkeit, einen Einblick in epistemologisches Denken über bestimmte Fächer zu erhalten.

Hofer führte eine Studie mit lautem Denken beim Prozess der Online-Suche nach Informationen über das Kommunikationsverhalten von Bienen durch. Den Versuchspersonen wurde die Aufgabe gegeben, innerhalb von 20 Minuten nach Informationen über das Kommunikationsverhalten von Bienen zu suchen, die zum Schreiben eines Kurzaufsatzes erforderlich sind und während der Informationssuche laut zu denken. Im Anschluss an die eigentliche Suche wurde ein retrospektives Interview zum Suchvorgang durchgeführt. Die Aufgabe wurde gewählt, weil hier insbesondere die Quelle und die Rechtfertigung der gefundenen Information beur-

tels einer Ratingskala beurteilt wird. Zur Bestimmung des Skalenwertes der Likert-Skala wird die Summe der Einzelbeurteilungen jedes Items berechnet (vgl. Bortz & Döring, 2006).

teilt werden müssen. Diese beiden Dimensionen sind nach Hofer durch Fragebogenverfahren nur problematisch zu erfassen.

Die Auswertung der Protokolle erfolgte nach den Vorgaben von Ericsson und Simon (1993) und es ergaben sich Hinweise auf alle vier Dimensionen des Modells von Hofer und Pintrich (1997). Die Äußerungen der Versuchspersonen beim lauten Denken ließen sich auf die metakognitive Beurteilung der Quellen sowie deren Rechtfertigung zurückführen. Allerdings bedeutet dieses Ergebnis nicht, dass die Versuchspersonen sich dieser epistemologischen Beurteilung bewusst sind. Die Quelle des Wissens war dabei schwieriger aus den Äußerungen abzuleiten als Prozesse zur Überwachung der Rechtfertigung. Diese metakognitiven Prozesse interagieren und scheinen mit motivationalen Prozessen verbunden zu sein. Beispielsweise wiesen Versuchspersonen mit der Überzeugung, dass Wissen einfach und sicher ist, nur eine geringe Tendenz auf, nach weitergehenden Informationen zu suchen. Weiterhin wurde von diesen Probanden auch wenig über die Glaubwürdigkeit der Quelle nachgedacht. Damit wird auch der Zusammenhang zwischen epistemologischen Überzeugungen und Motivation aus dem Wirkmodell von Hofer (2001) bestätigt. Weiterhin ergab sich ein ausgeprägter Zusammenhang mit der Expertise der Versuchspersonen. Weiter fortgeschrittene Versuchspersonen (hier Studenten) zeigten weniger naive epistemologische Überzeugungen. Damit einher ging eine stärkere Überwachung der Wissensquelle und Rechtfertigung. Ein letztes Ergebnis ist die Domänenspezifität von epistemologischen Überzeugungen. Es zeigte sich, dass die Beurteilung epistemologischer Aspekte in einer Disziplin nicht auf eine andere Disziplin übertragbar ist. So war in der gegebenen Untersuchung ein Teil der Versuchspersonen im Hauptfach Naturwissenschaftler, ein anderer Teil bestand aus Nicht-Naturwissenschaftlern. Nur die Naturwissenschaftler benutzen bei der Suche entsprechende Literaturdatenbanken, während sich die Nicht-Naturwissenschaftler herkömmliche Suchmaschinen zur Recherche zunutze machten. Die Überwachung der Wissensquelle war also abhängig von der Domäne. Aus einer normativen Perspektive sprechen diese Befunde dafür, dass Nicht-Naturwissenschaftler sich auf naivere epistemologische Überzeugung stützten.

Insgesamt eröffnet die Betrachtung epistemologischer Überzeugungen als Teil der metakognitiven Kapazitäten eines Individuums neue theoretische sowie methodische Möglichkeiten. Aus theoretischer Sicht ist die kognitive Betrachtung epistemologischer Überzeugungen interessant, da sie eine detailliertere Beschreibung der im Individuum ablaufenden Prozesse erlaubt, als dies mit dem Persönlichkeitstheoretischen Ansatz von Schommer (1990) bzw. dem

Modell epistemologischer Überzeugungen als Persönliche Theorien (Hofer & Pintrich, 1997) möglich ist. Daraus ergeben sich auch andere methodische Ansätze als die üblicherweise verwendeten Fragebogen und Interviewverfahren. Es rücken Verfahren in den Vordergrund, welche es gestatten, kognitive Prozesse zu erfassen. Eines der prominentesten Verfahren ist das zuvor im Beispiel vorgestellte laute Denken. Aber auch Verfahren, die Verhalten erfassen, können eingesetzt werden. So wurde im erwähnten Beispiel erfasst, welche Art Suchmaschine von Naturwissenschaftlern und Nicht-Naturwissenschaftler gewählt wurde, um daraus einen Rückschluss auf den abgelaufenen Prozess zu erhalten.

1.6 Pädagogische Relevanz epistemologischer Überzeugungen

Nach der Erläuterung einiger Modellvorstellungen epistemologischer Überzeugungen muss deren Einfluss auf pädagogisch relevante Konstrukte noch aufgezeigt werden. Die Pädagogische Psychologie als Anwendungsfach konzentriert sich auf die Erklärung, Beschreibung und Vorhersage lernrelevanten Verhaltens durch psychologische Theorien.²⁴ In diesem Kapitel soll dementsprechend die Relevanz epistemologischer Überzeugungen für das Lernen von Individuen beschrieben werden. Zur Beschreibung der Beziehungen zwischen epistemologischen Überzeugungen und Lernen wird hauptsächlich auf zwei bereits aufgeführte Modellvorstellungen zurückgegriffen, und zwar auf das Modell von Hofer (2001) und das EBS von Schommer-Aikins (2004). Zuerst werden beide Modelle im Hinblick auf den Einfluss epistemologischer Überzeugungen auf Lernen bzw. auf Lernen bezogenes Verhalten charakterisiert und anschließend die für die vorliegende Arbeit relevanten Aspekte herausgearbeitet. Ein Schwerpunkt wird dabei die Beziehung epistemologischer Überzeugungen zu Lernstrategien sein. Aus diesem Grund schließt dieser Kapitel mit einer Darstellung des Verfahrens, welches in den durchgeführten empirischen Untersuchungen zur Erfassung von Lernstrategien eingesetzt wurde.

1.6.1 Zusammenhang epistemologischer Überzeugungen und Lernstrategien

Schommer-Aikins (2004) spezifiziert in dem von ihr erarbeiteten EBS den Zusammenhang zwischen epistemologischen Überzeugungen, Lernstrategien und den eigentlichen Lernleistungen (vgl. Abbildung 8). Demnach beeinflussen Überzeugungen zu Lernen, Wissen und Wissenserwerb die Lernleistung und das selbstregulierte Lernen. Schommer nimmt hierbei

²⁴ An dieser Stelle soll keine Definition Pädagogischer Psychologie versucht werden. Es wird lediglich die Pädagogische Psychologie im Sinne der allgemeinen Definition der Wissenschaft Psychologie als Beschreibung, Erklärung und Vorhersage menschlichen Verhaltens im Kontext von Lehren und Lernen charakterisiert (vgl. Zimbardo, 1995).

teilweise eine direkte Wirkung epistemologischer Überzeugungen und eine über die Lernstrategien vermittelte Wirkung auf den Lernerfolg an. Im Gegensatz zu Schommer-Aikins (2004) geht Hofer (2001) nur von einer indirekten Wirkung epistemologischer Überzeugungen auf den Lernerfolg aus (vgl. Abbildung 6). In ihrem Modell wirken epistemologische Überzeugungen auf die Lernstrategien, die Lernmotivation sowie die Überzeugungen zum Unterricht und Lernen einer Person. Die Überzeugungen zum Lernen und Unterricht sowie die Lernmotivation wirken wiederum auf die Wahl der Lernstrategien. Die gewählten Lernstrategien und Lernmotivation bestimmen ihrerseits schließlich das Lernen, welches von Hofer als Wissenserwerb bzw. Wissensveränderung angesehen wird. Dies entspricht auch der Ansicht von Biggs (1993), wonach Lernstrategien vermittelnde Prozesse zwischen Lernen und dem Lernergebnis darstellen.

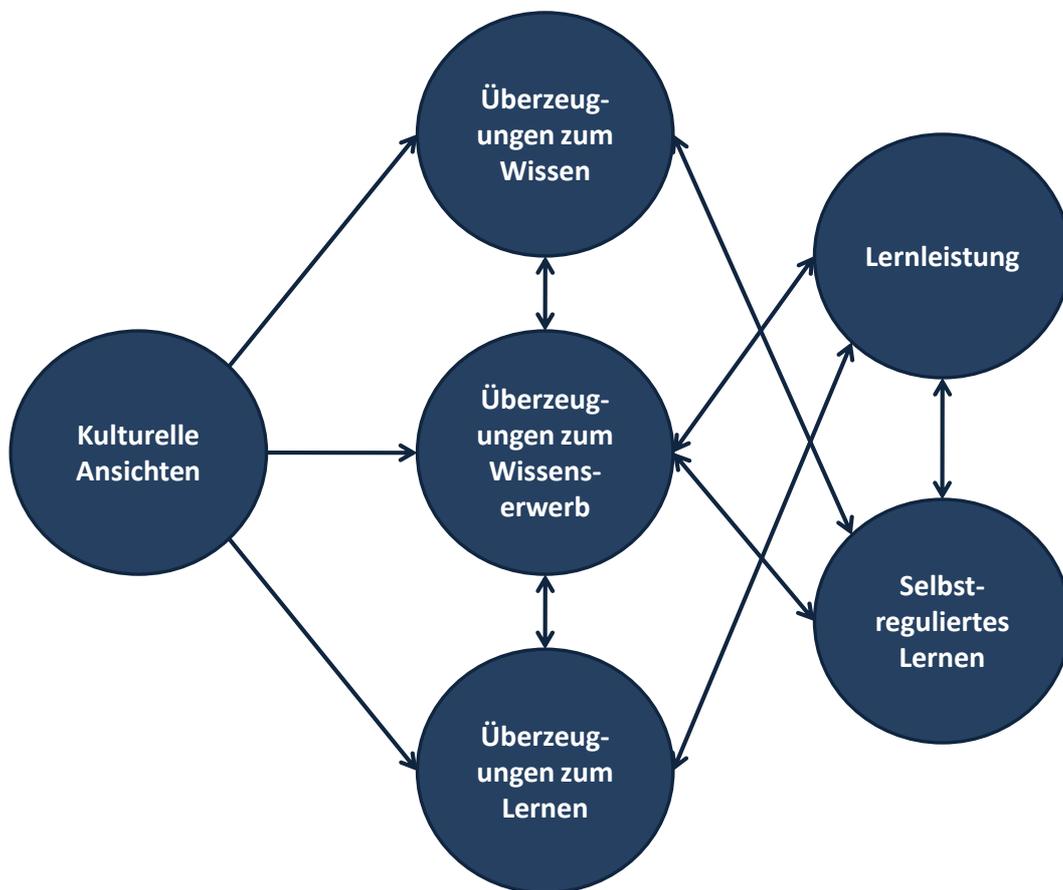


Abbildung 8: Das EBS von Schommer-Aikins

Eine Studie, die das Modell von Hofer (2001) aufgreift, wurde von Urhahne und Hopf (2004) durchgeführt. Dazu wurden die epistemologischen Überzeugungen mit einer von den Autoren übersetzten Version des Fragebogens von Conley et al. (2004) erfasst. Dieser Fragebogen beinhaltet die Dimensionen *Quelle*, *Sicherheit*, *Entwicklung* und *Rechtfertigung des Wissens*

und bezieht sich explizit auf epistemologische Überzeugungen in Bezug auf die Naturwissenschaften. Es wurde der Teil des Modells von Hofer geprüft, der die Beziehungen epistemologischer Überzeugungen zu Lernmotivation, Lernstrategien und den Überzeugungen über Lernen und Unterricht thematisiert. Aus dem Bereich der Lernmotivation wurden die Leistungsmotivation nach Heckhausen (1989) und das Interesse an Naturwissenschaften nach Krapp (2001) erhoben. Aus dem Bereich der Lernstrategien wurden die Lernstrategien nach Wild und Schiefele (1994) erfasst, wobei hier nur die Wiederholungs-, Elaborations-, Anstrengungs- und Kontrollstrategien berücksichtigt wurden (siehe Kapitel 1.6.2 für eine detaillierte Beschreibung dieser Lernstrategien).²⁵ Im Bereich der Überzeugungen zum Lernen und Unterricht wurden die fachspezifischen Selbstkonzepte für Biologie und Physik nach Moschner (2001) erhoben. Die Zusammenhänge zwischen epistemologischen Überzeugungen und Motivation, Lernstrategien und Selbstkonzept wurde an einer Stichprobe von 167 Schülern der neunten Jahrgangsstufen dreier Gymnasien erhoben. Davon waren 102 Jungen und 65 Mädchen mit einem mittleren Alter von 15.36 Jahren ($SD = .61$). Die erhobenen Variablen wurden mit den Skalen zur Erfassung der epistemologischen Überzeugungen korreliert, die Ergebnisse sind in Tabelle 22 dargestellt.

Tabelle 22: Ergebnisse von Urhahne und Hopf (2004)

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

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

Die Tabelle gibt die bivariaten Korrelationen von epistemologischen Überzeugungen und Lernstrategien wieder, die auf deutliche Zusammenhänge zwischen epistemologischen Überzeugungen und Lernstrategien hindeuten. Es lässt sich auch die Hypothese bestätigen, dass fortgeschrittene epistemologische Überzeugungen mit einer geringeren Verwendung von

²⁵ Diese Auswahl der erhobenen Lernstrategien orientiert sich an der Auswahl in der PISA-2000-Studie (Artelt, Baumert, Julius-McElvany, 2003).

Wiederholungsstrategien und einer häufigeren Verwendung von Elaborations-, Kontroll- und Anstrengungsstrategien einhergehen. Ersichtlich wird dies an den negativen Korrelationen von Wiederholungsstrategien mit epistemologischen Überzeugungen bzw. an den positiven Korrelationen von Elaborationsstrategien mit epistemologischen Überzeugungen. Allerdings gilt diese Schlussfolgerung nicht für die Dimension *Quelle des Wissens*. Bei dieser Dimension scheint die Überzeugung, dass Wissen durch Autoritäten wie Lehrer vermittelt wird, mit einer erhöhten Anstrengung und kognitiv anspruchsvolleren Lernstrategien einherzugehen. Die Ergebnisse von Urhahne und Hopf (2004) bestätigen einen Zusammenhang von Lernstrategien und epistemologischen Überzeugungen. Allerdings ist mit der Verwendung von bivariaten Korrelationen nicht mehr als eine deskriptive Beschreibung der Zusammenhänge möglich, da keine Wirkrichtung angegeben werden kann.

Cano (2005) hat ein Pfadmodell der Wirkung von epistemologischen Überzeugungen auf Lernstrategien und die Lernleistung aufgestellt. Er geht hierbei von einer direkten und indirekten Wirkung epistemologischer Überzeugungen auf die Lernleistungen aus (s. Abbildung 9). Das Modell spezifiziert direkte Wirkungen epistemologischer Überzeugungen auf die Lernleistung und eine indirekte Wirkung epistemologischer Überzeugungen, die über die Lernstrategien vermittelt werden.

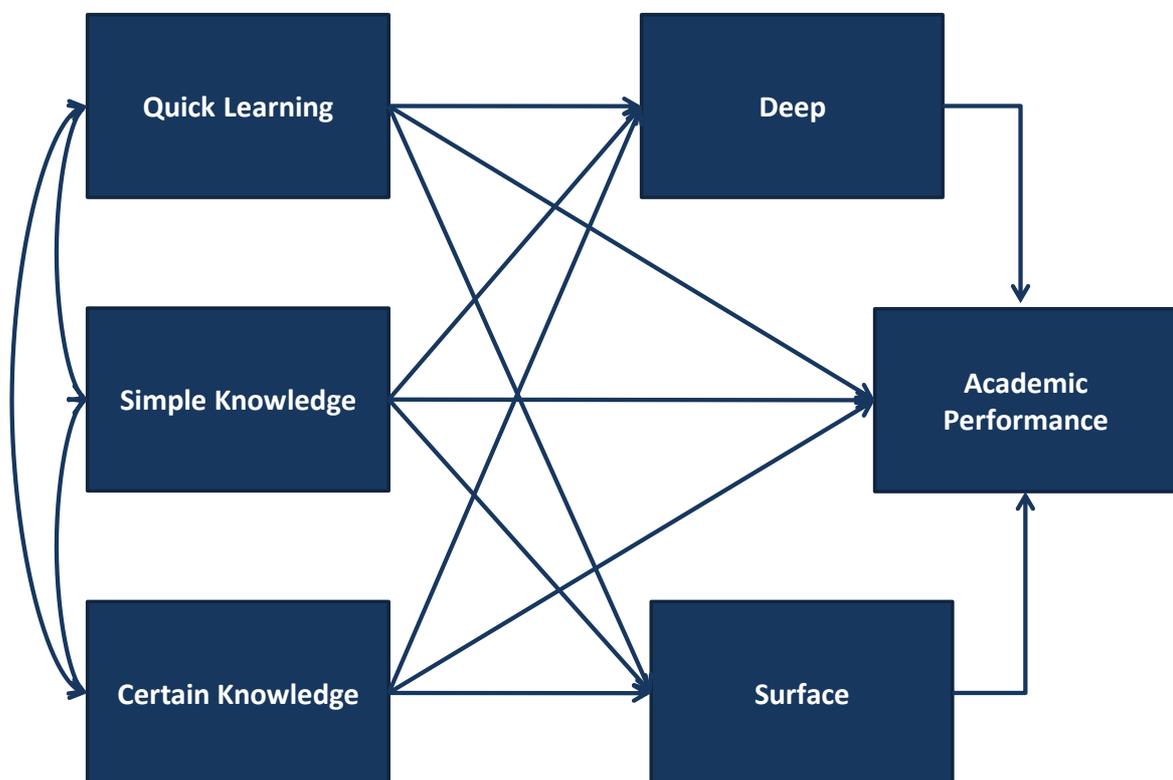


Abbildung 9: Pfadmodell von Cano (2005)

Zur Überprüfung dieses Modells wurde eine Untersuchung mit 1600 spanischen Schülern durchgeführt, davon 722 weibliche und 878 männliche. Das mittlere Alter der Versuchspersonen lag bei 14.79 Jahren ($SD = 1.95$). Epistemologische Überzeugungen wurden mithilfe des EBQ erfasst. Cano fand hier mithilfe einer explorativen Faktorenanalyse die drei Faktoren *Quick Learning*, *Simple Knowledge* und *Certain Knowledge*. Zur Erfassung der Lernstrategien wurde der *Learning Process Questionnaire* von Barca (1999) verwendet. Dieser Fragebogen ist eine spanische Version des *Study Process Questionnaire* (SPQ) von Biggs (1987) und soll zwei grundlegende Strategien des Lernens erfassen: zum einen die tiefenbezogene Strategie, bei welcher der Lerner versucht, die Intention eines Textes zu verstehen sowie die Lerninhalte mit dem Vorwissen und der eigenen persönlichen Erfahrung zu verbinden, und zum anderen die oberflächenbezogene Strategie, die lediglich dem Ziel dient, den Lerninhalt wiedergeben zu können. Der verwendete Fragebogen umfasst fünf Skalen: *surface strategy*, *surface motive*, *deep strategy*, *deep motive* und *achieving motive*. Das *achieving motive* ist eine Erweiterung des Ansatzes von Marton und Säljö (1976) von Biggs (1987) und gibt die Motivation an, gute Noten zu erreichen. Eine explorative Faktorenanalyse ergab zwei Dimensionen, wobei die Skalen *deep strategy*, *deep motive* und *achieving motive* auf dem ersten Faktor luden und die Skalen *surface strategy* und *surface motive* auf dem zweiten Faktor. Hieraus wurden die beiden Skalen *Deep* und *Surface* gebildet. Zur Bestimmung der Leistung der Schüler wurden am Schuljahresende die Noten erhoben. Die Durchschnittsnote wurde als Indikator für die schulische Leistung herangezogen. Das in Abbildung 9 dargestellte Pfadmodell zeigte eine gute Modellpassung, jedoch waren einige Pfade nicht signifikant, sodass sie sukzessive eliminiert wurden.

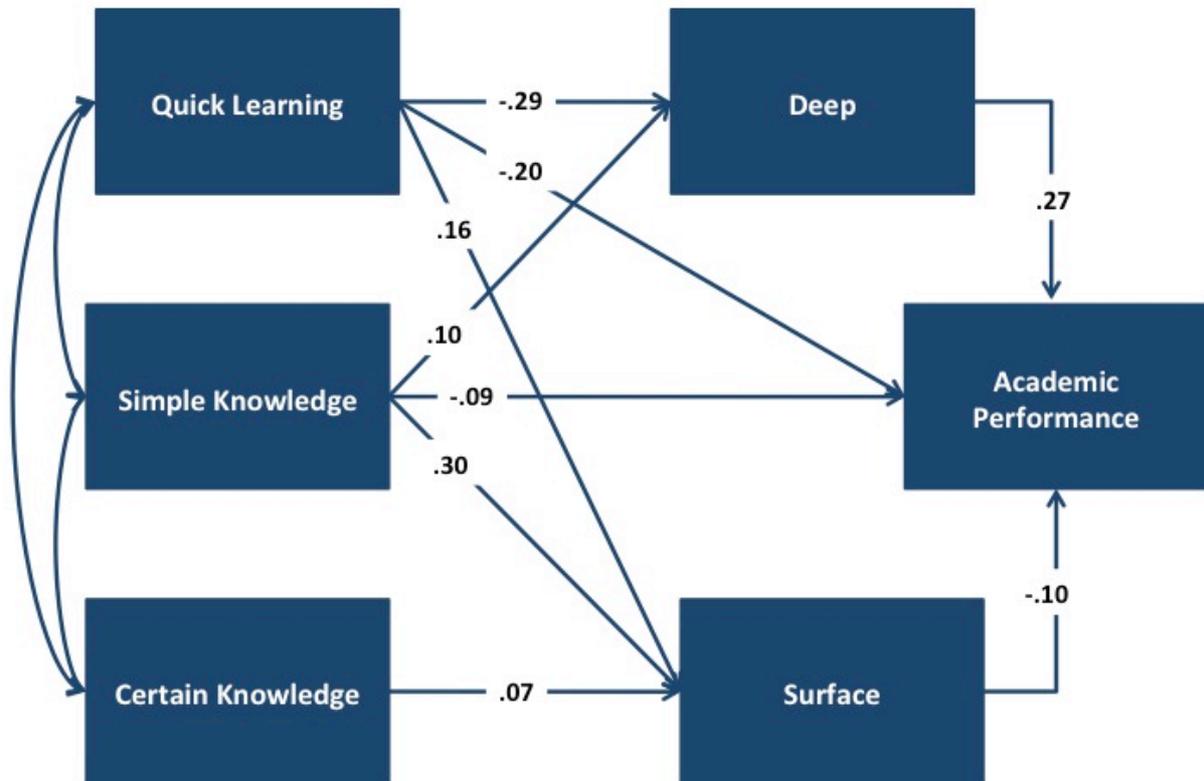


Abbildung 10: Resultierendes Pfadmodell mit signifikanten Pfaden nach Cano (2005)

Die verbleibenden signifikanten Pfade sind als standardisierte Koeffizienten in Abbildung 10 zu sehen. Die Fitindizes des Gesamtmodells waren $\chi^2 = 5.19$ ($df = 3$), $p = .16$, $GFI = 1.00$, $AGFI = .99$ und $RMR = .01$. Das Pfadmodell zeigt, dass die schulische Leistung nur von *Quick Learning* und *Simple Knowledge* direkt beeinflusst wurde. Beide Pfadkoeffizienten sind negativ, was unter der Berücksichtigung der Kodierung der Antwortskalen bedeutet, dass naivere epistemologische Überzeugungen in Hinblick auf einfach strukturiertes Wissen und die Schnelligkeit des Lernprozesses zu schlechteren schulischen Leistungen führen. *Quick Learning* übt weiterhin einen negativen Einfluss auf die Verwendung tiefenbezogener Lernstrategien aus, wobei der Einfluss von *Quick Learning* auf die Verwendung von oberflächenbezogenen Lernstrategien positiv war. *Simple Knowledge* beeinflusst sowohl die Verwendung tiefenbezogener als auch oberflächenbezogener Lernstrategien positiv. *Certain Knowledge* wirkte nur sehr schwach auf die Verwendung oberflächenbezogener Lernstrategien. Eine tiefenbezogene Lernstrategie beeinflusste den schulischen Erfolg positiv, während oberflächenbezogene Lernstrategien diesen negativ beeinflussten. Insgesamt ergab sich also ein Einfluss epistemologischer Überzeugungen auf die Verwendung von Lernstrategien und die schulische Leistung, wobei jedoch nicht alle epistemologischen Überzeugungen direkt, sondern nur vermittelt über die Lernstrategien auf die schulische Leistung wirkten.

Aufbauend auf der Arbeit von Cano (2005) hat Phan (2006) den Zusammenhang von epistemologischen Überzeugungen, Lernstrategien, reflexivem Denken und Lernleistung untersucht.

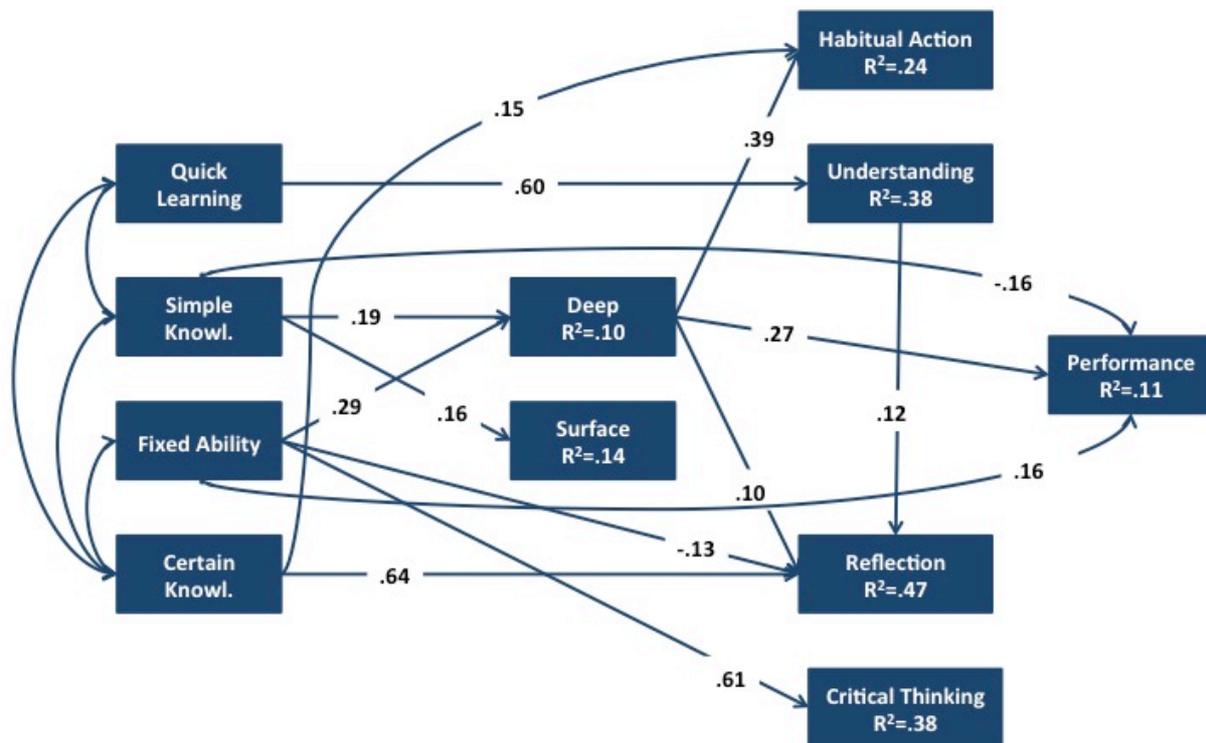


Abbildung 11: Pfadmodell von Phan (2006)

In diesem Pfadmodell wird davon ausgegangen, dass epistemologische Überzeugungen sowohl direkt auf die Lernleistung als auch indirekt über die angewandten Lernstrategien und reflexives Denken auf die Lernleistung wirken. In Erweiterung des Modells von Cano kommt hier somit das reflexive Denken als Mediator hinzu. Die Erfassung epistemologischer Überzeugungen erfolgte auch hier mit dem EBQ, die Lernstrategien wurden mit dem revidierten *2-Factors-Study-Progress-Questionnaire* (R-SPQ-2F) von Biggs, Kember und Leung (2001) erhoben. Der R-SPQ-2F erfasst die Faktoren *surface strategy*, *surface motive*, *deep strategy* und *deep motive*. Das reflexive Denken wurde mithilfe des *Reflective Thinking Questionnaire* (RTQ) von Kember et al. (2000) festgestellt. Dieser Fragebogen erfasst die vier Stufen reflexiven Denkens nach Mezirow (1991, 1998), der *Habitual action*, *Understanding*, *Reflection* und *Critical Reflection* unterscheidet. *Habitual action* meint die einfache und fast automatische Informationsaufnahme, ohne das Gelernte bewusst zu verarbeiten. *Understanding* bedeutet die Verarbeitung des Gelernten, aber ohne eine Verbindung zu anderen Wissensinhalten herzustellen. *Reflection* bedeutet die aktive und sorgfältige Beachtung aller Annahmen, die dem Gelernten zugrunde liegen. *Critical Reflection* bildet eine fortgeschrittene Stufe von *Reflection* und erweitert diese um das Bewusstsein, warum Dinge auf eine bestimmte Art und

Weise wahrgenommen werden und warum auf bestimmte Art und Weise gehandelt und gefühlt wird. Die Lernleistung wurde als Summe der Leistungen des am Semesterende erfolgreich abgeschlossenen Kurses ermittelt.

Die Prüfung des Pfadmodells aus Abbildung 11 wurde mit den Daten von 332 studentischen Versuchspersonen durchgeführt, davon 146 Frauen und 185 Männer mit einem durchschnittlichen Alter von 22 Jahren ($SD = 1.12$). Fast alle Versuchspersonen waren Studenten der Pädagogik. Zur Auswertung wurde der EBQ einer explorativen Faktorenanalyse unterzogen, die zu vier Faktoren führte. Dies waren *Quick learning*, *Simple knowledge*, *Certain knowledge* und *Fixed ability*. Die Items des R-SPQ-2F und RTQ wurden den Vorgaben entsprechend zu Skalen summiert. Die anschließende Prüfung des Pfadmodells ergab eine gute Modellpassung, jedoch wurden aufgrund der Modifikationsindizes die nicht signifikanten Pfade aus dem Modell entfernt. Das resultierende Modell (vgl. Abbildung 11) wies folgende Fitindizes auf: $\chi^2 = 39.64$, $df = 12$, $p < .01$, $CFI = .98$, $NNFI = .92$ und $RMSEA = .08$. Im Zusammenhang mit den verwendeten Lernstrategien ist in diesem Pfadmodell auffällig, dass nur die beiden Faktoren *Simple Knowledge* und *Fixed Ability* direkt auf die beiden Faktoren *Deep* und *Surface* wirken, ebenso wirken diese Faktoren direkt auf die am Ende des Semesters erhobene Lernleistung. Der Faktor *Fixed Ability* wirkt auch direkt auf die beiden Stufen *Reflection* und *Critical Thinking*. Die beiden Faktoren *Quick Learning* und *Certain Knowledge* wirken nicht auf die Wahl der Lernstrategien, sondern nur auf die Stufen des reflexiven Denkens. In dem Modell zeigte sich demnach auch ein klarer Zusammenhang zwischen epistemologischen Überzeugungen und Lernstrategien.

Phan (2008) untersuchte den Zusammenhang zwischen epistemologischen Überzeugungen, Lernstrategien und selbstgesteuertem Lernen mithilfe von multiplen Regressionsanalysen. Dabei wurden die epistemologischen Überzeugungen erneut mit dem EBQ und die Lernstrategien mit dem R-SPQ-2F von Biggs et al. (2001) erfasst. Die Strategien selbstgesteuerten Lernens wurden mithilfe der entsprechenden Skala aus dem *Motivated Strategies for Learning Questionnaire* (MSLQ) von Pintrich, Smith, Garcia und MacKeachie (1991) erhoben. Im Gegensatz zu der vorgestellten Untersuchung von Phan (2006) waren in dieser Studie sowohl die epistemologischen Überzeugungen als auch die Lernstrategien die Prädiktoren und die Strategien selbstgesteuerten Lernens bildeten das Kriterium der Regressionsanalyse. Die Untersuchung wurde an 603 Versuchspersonen, davon 266 Frauen und 337 Männer, durchgeführt, die ein Durchschnittsalter von 25.2 Jahren ($SD = 1.49$) hatten und Studenten im ersten

Semester waren. Aus dem Fragebogen von Schommer wurden mittels einer explorativen Faktorenanalyse vier Faktoren extrahiert: *Fixed ability*, *Certain knowledge*, *Simple knowledge* und *Quick learning*. Aus dem R-SPQ-2F wurden die Skalen für *Surface strategy*, *Surface motive*, *Deep strategy* und *Deep motive* gebildet. Zusätzlich wurde das Geschlecht als Prädiktor in die multiple Regression aufgenommen. Das Kriterium dieser Regressionsanalyse stellte das selbstgesteuerte Lernen dar, welches mit der entsprechenden Skala des MSLQ erhoben wurde. Die Ergebnisse sind in Tabelle 23 dargestellt, aufgeführt werden die standardisierten und nicht standardisierten Regressionskoeffizienten.

Tabelle 23: Ergebnisse von Phan (2008)

Variable	b-Koeffizient	β -Koeffizient
Gender	-.03	-.01
Deep motive	.12	.19***
Deep strategy	.04	.05
Surface motive	-.07	-.01
Surface strategy	.01	.02
Fixed ability	-.10	-.11*
Certain knowledge	-.01	.01
Simple knowledge	-.09	-.10
Quick learning	-.05	-.06

Anmerkung: *: $p < .05$, ***: $p < .001$

Die Varianzaufklärung dieses Regressionsmodells betrug 13.2 %, das Modell ist signifikant mit $F(9, 593) = 9.99, p < .001$. *Fixed Ability*, *Simple Knowledge* sowie *Deep motive* erwiesen sich als signifikante Prädiktoren.

In den letzten Kapiteln wurden nur Untersuchungen betrachtet, die ein mehr oder weniger enges Spektrum möglicher Lernstrategien analysiert haben. Dies waren tiefen- und oberflächenbezogene Lernstrategien nach Marton und Säljö (1976). Neben diesen beiden Kategorien von Lernstrategien gibt es noch weitergehende Taxonomien. Die Verwendung einer feineren Unterteilung von Lernstrategien erlaubt es, den Einfluss epistemologischer Überzeugungen auf Lernstrategien genauer zu untersuchen.

Dahl, Bals und Turi (2005) führten eine Untersuchung durch, bei der alle Skalen des MSLQ von Pintrich et al. (1991) benutzt wurden. Der MSLQ erfasst fünf verschiedene Arten von Lernstrategien, darunter fallen Wiederholungsstrategien, Elaborierungsstrategien, Strategien des kritischen Denkens, Organisationsstrategien sowie metakognitive und selbstregulative Strategien. Der EBQ wurde zur Erfassung epistemologischer Überzeugungen verwendet. Ziel der Studie war es, die Beziehung zwischen epistemologischen Überzeugungen und den mit-

hilfe des MSLQ erfassten Lernstrategien zu ermitteln. Die epistemologischen Überzeugungen wurden hierzu als Prädiktoren einer Regressionsanalyse aufgenommen. Die Lernstrategien stellten jeweils die Kriterien dar. Die Untersuchung wurde mit 81 studentischen Versuchspersonen durchgeführt, davon 21 Männer und 60 Frauen mit einem Durchschnittsalter von 22.6 Jahren ($SD = 6.4$). Zur Bestimmung der epistemologischen Überzeugungen *Simple Knowledge*, *Certain Knowledge*, *Fixed Ability* und *Quick Learning* wurden die Items entsprechend der von Schommer (1995) vorgegebenen Prozedur zusammengefasst. Hierbei werden die Items dem Auswertungsschlüssel entsprechend zu Skalen addiert, die Skalenwerte anschließend z-standardisiert und mit den (von Schommer erhältlichen und unpublizierten) Faktorkoeffizienten multipliziert. Die Skalen des MSLQ wurden gemäß der Zuordnung der Items zu den fünf erfassten Lernstrategien gebildet. Die Interkorrelation der Skalen ist in Tabelle 24 aufgeführt. Die Korrelationsmatrix deutet erstens auf eine Korrelation der Dimensionen epistemologischer Überzeugungen hin, ebenso korrelieren die Lernstrategien des MSLQ untereinander sehr hoch. Zweitens zeigte sich, dass vor allem die Dimensionen *Simple Knowledge* und *Fixed Ability* mit den Lernstrategien korrelieren. Die Dimension *Simple Knowledge* zeigte keine signifikante Korrelation mit kritischem Denken und ebenso korrelierte die Dimension *Fixed Ability* nicht mit Wiederholungsstrategien. Von den Dimensionen *Quick Learning* und *Certain Knowledge* korrelierte lediglich *Certain Knowledge* signifikant mit den Organisationsstrategien. Bei den nachfolgenden multiplen Regressionsanalysen ist also zu erwarten, dass vor allem die Dimensionen *Simple Knowledge* und *Fixed Ability* bedeutsame Prädiktoren darstellen.

Tabelle 24: Interkorrelation von epistemologischen Überzeugungen und Lernstrategien (Dahl et al., 2005)

	1.	2.	3.	4.	5.	6.	7.	8.
1. Simple	-							
2. Fixed	.29**	-						
3. Quick	.18	-.27*	-					
4. Certain	.06	-.37**	.03	-				
5. Wiederholungsstrategien	-.31**	-.16	.03	-.09	-			
6. Elaborationsstrategien	-.27*	-.44**	-.16	-.13	.41**	-		
7. Kritisches Denken	-.15	-.50**	-.13	-.17	.26*	.59 **	-	
8. Organisationsstrategien	-.43**	-.32**	-.04	-.22*	.55**	.62 **	.42 **	-
9. Metakognitive und selbstregulative Strategien	-.41**	-.35**	-.08	-.15	.56**	.62 **	.53 **	.66 **

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

Die multiplen Regressionsanalysen sind in Tabelle 25 zusammengefasst. Prädiktoren waren die Dimensionen epistemologischer Überzeugungen. Die Kriterien waren jeweils die Skalen des MSLQ. Dargestellt werden jeweils die standardisierten Regressionskoeffizienten, die t-Werte sowie die durch das Regressionsmodell aufgeklärte Varianz und die Gesamtsignifikanz des Modells.

Tabelle 25: Multiple Regressionen epistemologischer Überzeugungen auf Lernstrategien (Dahl et al., 2005)

Prädiktor	Wiederholungsstrategien		Elaborationsstrategien		Kritisches Denken		Organisationsstrategien		Metakognitive/selbstregulative Strategien	
	β	t	β	t	β	t	β	t	β	t
Simple	-.30	-2.60**	-.15	-1.43	.00	.00	-.38	-3.64**	-.34	-3.21**
Fixed	-.08	-.64	-.39	-3.37***	-.50	-4.37***	-.18	-1.52	-.25	-2.16*
Quick	.10	.89	-.03	-.27	.00	.04	.08	.79	.05	.46
Certain	-.04	-.36	.03	.28	.02	.16	-.14	-12.5	-.04	-.33
R ²	.11		.22		.25		.25		.23	
F-Wert	2.33		5.24***		6.22***		6.16***		5.65***	

Anmerkung: **: $p < .05$, ***: $p < .001$

Bei den Wiederholungsstrategien ist lediglich *Simple Knowledge* ein signifikanter Prädiktor. Allerdings ist das Gesamtmodell selbst nicht signifikant und die Varianzaufklärung beträgt lediglich 11 %, dieses Modell klärt von allen Regressionsanteilen die geringste Varianz auf. Bei den Elaborationsstrategien ist nur *Fixed Ability* ein signifikanter Prädiktor. Das gesamte Regressionsmodell klärt 22 % der Varianz auf und ist signifikant. Beim kritischen Denken ist *Fixed Ability* ebenso der einzig signifikante Prädiktor, dieses Modell klärt 25 % der Varianz auf und ist signifikant. Auch bei den Organisationsstrategien ist nur *Simple Knowledge* ein signifikanter Prädiktor. Dieses Modell klärt ebenfalls 25 % Varianz auf und ist signifikant. Bei den metakognitiven und selbstregulativen Strategien sind *Simple Knowledge* und *Fixed Ability* signifikante Prädiktoren, dieses Modell ist signifikant und klärt 23 % der Varianz auf.

Es zeigt sich, dass epistemologische Überzeugungen bei der Vorhersage von Lernverhalten nur bedingt nützlich sind, da die Varianzaufklärung der Modelle nicht sehr groß ist. Dieses Ergebnis entspricht dem der Studie von Phan (2006), bei der sich ebenfalls eine eher geringe Varianzaufklärung ergab. Überdies sind nur zwei der vier epistemologischen Überzeugungen überhaupt bedeutsame Prädiktoren. Dahl et al. (2005) schließen daraus, dass nur einige der von Schommer (1990) konzeptualisierten epistemologischen Überzeugungen für die Vorhersage des selbst berichteten Gebrauchs verschiedener Lernstrategien geeignet sind. Darüber hinaus deutet die geringe Varianzaufklärung der Regressionsmodelle darauf hin, dass die

epistemologischen Überzeugungen nur zu einem kleinen Teil die Wahl von Lernstrategien beeinflussen. Allerdings steht einer Verallgemeinerung dieser Ergebnisse die geringe Stichprobengröße von nur 81 Probanden entgegen, was zu einer ineffizienten Schätzung der Parameter führen kann.

Zusammenfassend lässt sich aus den Studien von Cano (2005), Phan (2006, 2008) und Dahl et al. (2005) dennoch schließen, dass epistemologische Überzeugungen die Wahl von Lernstrategien beeinflussen, auch wenn die Effekte im geringen bis mittleren Bereich angesiedelt sind und noch andere Variablen existieren, welche einen Einfluss auf die Verwendung von Lernstrategien ausüben. Die Existenz von anderen, nicht in den Regressionsmodellen vorkommenden Variablen wird durch deren geringe Varianzaufklärung nahegelegt. Die aufgeführten Studien konnten aber dennoch belegen, dass epistemologische Überzeugungen die Lernstrategien von Personen beeinflussen, unabhängig davon, ob noch andere Variablen eine Rolle spielen können.

1.6.2 LIST-Fragebogen zur Erfassung von Lernstrategien im Studium

Im Folgenden wird im Vorgriff auf die Forschungsfragen und die Darstellung der Methoden der LIST-Fragebogen von Wild und Schiefele (1994) beschrieben. Mithilfe des LIST-Fragebogens sollen Lernstrategien erfasst werden, um anschließend die Beziehungen zu epistemologischen Überzeugungen, welche wiederum mithilfe des EPI-Fragebogens von Moschner und Gruber erfasst werden, zu untersuchen.

Der LIST wurde ausgehend von dem zuvor kurz dargestellten MSLQ von Pintrich et al. (1991) sowie dem *Learning and Study Strategies Inventory* (LASSI) von Weinstein (1988) konzipiert und übernimmt deren Trennung von kognitiven und motivationalen Aspekten des Lernverhaltens. Allerdings konzentriert sich der LIST nur auf die kognitiven Aspekte. Vom MSLQ wurde für die Entwicklung des LIST die zweistufige Klassifikation der kognitiven Lernstrategien übernommen. Auf der ersten Stufe wird zwischen kognitiven, metakognitiven und ressourcenbezogenen Strategien unterschieden. Auf der zweiten Gliederungsstufe werden für diese drei Strategiearten Unterstrukturen festgelegt, die den Ausgangspunkt für die Konstruktion der Skalen bilden.

Kognitive Lernstrategien umfassen all jene Prozesse, die unmittelbar zur Informationsaufnahme, -verarbeitung und -speicherung dienen. Unter die kognitiven Lernstrategien fallen

insbesondere die Wiederholungs-, Organisations- und Elaborationsstrategien. Metakognitive Strategien beziehen sich auf die aktiven und bewussten Kontroll- und Steuerungsprozesse, die während des Lernens benutzt werden. Die metakognitiven Strategien Planen, Überwachen und Regulieren bilden somit einen Regelkreis ab, der bei der Planung der Lerntätigkeit beginnt, zur Überwachung der Lerntätigkeit übergeht und im Anschluss an die Lerntätigkeit eine Überprüfung des selbst geplanten Lernziels durchführt und nachfolgende Lerntätigkeiten reguliert. Ressourcenbezogene Lernstrategien sind Aktivitäten, die der Lerner durchführt, um die Lerntätigkeit als solche zu organisieren. Hierbei wird wiederum nach internen und externen Ressourcen unterschieden. Interne Ressourcen betreffen die Allokation eigener Anstrengung, Aufmerksamkeit, Konzentration und Zeit, externe Ressourcen betreffen die Nutzung externer Informationsquellen, wie bspw. Literatur oder kooperativen Lernens (Wild, 2005).

Ausgehend von dieser Konzeption wurde von Wild und Schiefele (1994) eine Sammlung von Items aufgestellt, die der Erfassung dieser Lernstrategien dienen sollte. Dabei orientierten sich die Autoren am MSLQ und LASSI sowie dem SPQ von Biggs (1987). Zusätzlich wurden Items durch die Befragung von Studenten verschiedener Fachrichtung erweitert und überarbeitet. Diese Items wurden anschließend an 310 Studenten aus den Bereichen Bau- und Vermessungswesen, Elektrotechnik, Informatik, Luft- und Raumfahrttechnik sowie Wirtschaftswissenschaften verteilt. Die Auswertung erfolgte durch eine Hauptkomponentenanalyse, welche für jeden der drei Teilbereiche der Lernstrategien getrennt durchgeführt wurde.

In Tabelle 26 sind die Ergebnisse der Studien von Wild und Schiefele (1994) dargestellt. Die Beschreibung der Skalen sowie die dazugehörigen konkreten Verhaltensweisen sind Wild, Schiefele und Winteler (2009) entnommen. Im Bereich der kognitiven Lernstrategien ergab sich eine 4-Komponentenlösung. Aus diesen vier Komponenten wurden die Skalen *Elaborieren*, *Kritisches Prüfen*, *Organisieren* und *Wiederholen* gebildet. Im Bereich der metakognitiven Strategien konnte bei einer 3-Komponentenlösung keine eindeutige Zuordnung der Items zu den Hauptkomponenten vorgenommen werden, sodass eine Skala *Metakognitive Strategien* mit allen Items, die das Planen, Überwachen und Regulieren von Lerntätigkeiten beschreiben, gebildet wurde. Im Bereich der ressourcenorientierten Strategien wurde die Analyse auf die Bereiche externe und interne Ressourcen aufgeteilt. Im Bereich der internen Ressourcen ergab sich eine 3-Komponentenlösung und die Items wurden zu Skalen zusammengefasst, welche die Strategien *Anstrengung*, *Aufmerksamkeit* und *Zeitmanagement* erfassen. Im Bereich der externen Ressourcen ergab sich ebenfalls eine 3-Komponentenlösung und die

Items wurden zu Skalen zusammengefasst, welche die Strategien *Lernumgebung*, *Lernen mit Studienkollegen* und *Literatur* erfassen. Die Reliabilitäten der gebildeten Skalen sind ebenfalls in Tabelle 26 aufgeführt. Am Ende der Skalenbildung umfasste der komplette LIST-Fragebogen 77 Items.

Tabelle 26: Übersicht über die Lernstrategien nach Wild et al. (2009)

Skala/Cronbachs α	Beschreibung
Kognitive Lernstrategien	
Elaborieren $\alpha = .82$	Die Skala erfasst Studientätigkeiten, die auf ein tieferes Verstehen des Stoffes ausgerichtet sind. Dies geschieht, indem neuer Stoff in ein Netzwerk anderer Bezüge eingebettet wird (z. B. 'Ich versuche, neue Begriffe oder Theorien auf mir bereits bekannte Begriffe und Theorien zu beziehen.')
Kritisches Prüfen $\alpha = .72$	Die Skala erfasst Studientätigkeiten, die das Verständnis für den Stoff durch ein kritisches Hinterfragen von Aussagen und Begründungszusammenhängen vertiefen (z. B. 'Ich prüfe, ob die in einem Text (oder in meiner Mitschrift) dargestellten Theorien, Interpretationen oder Schlussfolgerungen ausreichend belegt und begründet sind.').
Organisieren $\alpha = .77$	Die Skala erfasst Studientätigkeiten, die durchgeführt werden, um einen zu bewältigenden Stoff in geeigneter Weise zu reorganisieren (z. B. 'Ich fertige Tabellen, Diagramme oder Schaubilder an, um den Stoff der Veranstaltung besser strukturiert vorliegen zu haben.').
Wiederholen $\alpha = .73$	Die Skala erfasst Studientätigkeiten, die auf das Einprägen von Fakten und Regeln durch schlichtes Wiederholen (z. B. 'Ich lese meine Aufzeichnungen mehrmals hintereinander durch.')
Metakognitive Lernstrategien $\alpha = .64$	
Planung, Überwachung, Regulation	In einer Skala zusammengefasst werden die drei Teilaspekte 'Planung', 'Überwachung' und 'Regulation' erhoben (z. B. 'Ich versuche, mir vorher genau zu überlegen, welche Teile eines bestimmten Themengebiets ich lernen muss und welche nicht. '; 'Um Wissenslücken festzustellen, rekapituliere ich die wichtigsten Inhalte, ohne meine Unterlagen zu Hilfe zu nehmen.').
Ressourcenbezogene Strategien (interne Ressourcen)	
Anstrengung $\alpha = .74$	Die Skala erfasst, inwieweit vermehrte Anstrengungen in Kauf genommen werden, um Studien- und Lernziele zu erreichen (z. B. 'Ich lerne auch spätabends und am Wochenende, wenn es sein muss.').
Aufmerksamkeit $\alpha = .90$	Die Skala erfasst subjektiv wahrgenommene Aufmerksamkeitsfluktuationen (z. B. 'Wenn ich lerne, bin ich leicht abzulenken.').
Zeitmanagement $\alpha = .83$	Die Skala erfasst, inwieweit eine Zeitplanung vorgenommen und eingehalten wird (z. B. 'Ich lege bestimmte Zeiten fest, zu denen ich dann lerne.').

Skala/Cronbachs α	Beschreibung
Ressourcenbezogene Strategien (externe Ressourcen)	
Lernumgebung $\alpha = .71$	Die Skala erfasst, inwieweit eine äußere Lernumgebung geschaffen oder gesucht wird, die ein konzentriertes und ungestörtes Arbeiten ermöglicht (z. B. 'Die wichtigsten Unterlagen habe ich an meinem Arbeitsplatz griffbereit.').
Lernen mit Studienkollegen $\alpha = .82$	Die Skala erfasst das Ausmaß kooperativen Lernens. Sie umfasst zum einen verschiedene Formen gemeinsamer Arbeit; sie umfasst aber auch Formen einseitiger Inanspruchnahme von Studienkollegen (z. B. 'Ich bearbeite Texte oder Aufgaben zusammen mit meinen Studienkollegen.').
Literatur $\alpha = .72$	Die Skala erfasst, inwieweit auf zusätzliche Literatur zurückgegriffen wird (z. B. 'Fehlende Informationen suche ich mir aus verschiedenen Quellen zusammen').

Die Ergebnisse von Wild und Schiefele (1994) wurden von Boerner, Seeber, Keller und Beinborn (2005) validiert. Dabei wurde die faktorielle Validität untersucht. Zur Validierung wurden einige Items auf die Belange von berufstätigen Studierenden umformuliert. Insgesamt 20 Items wurden nach einer Vorstudie zur Skala zur Erfassung metakognitiver Strategien ersetzt bzw. hinzugefügt, sodass diese Version des LIST insgesamt 85 Items enthielt. Letzteres hatte vor allem das Ziel, die internen Konsistenzen dieser Skalen zu verbessern sowie ihre dreidimensionale Erfassung zu gewährleisten. Der Fragebogen wurde schließlich von insgesamt 577 Studierenden bearbeitet.

Die Prüfung der faktoriellen Validität erfolgte in zwei Stufen. In der ersten wurden alle Items einer Hauptkomponentenanalyse unterzogen. Es wurden 13 Hauptkomponenten extrahiert und einer Varimax-Rotation unterzogen. Auf diesen Hauptkomponenten wurden alle Items mit Doppelladungen und Ladungen kleiner als .50 entfernt. Anschließend wurden die verbleibenden Items einer erneuten Hauptkomponentenanalyse mit der Extraktion von 13 Faktoren und Varimax-Rotation unterzogen. Auf diese Weise konnten bis auf die Strategie *Wiederholen* alle Strategien von Wild und Schiefele (1994) repliziert werden. Überdies zeigte sich für die metakognitiven Strategien die geforderte dreidimensionale Struktur mit den Strategien *Planung*, *Kontrolle* und *Regulation*. Somit konnte die faktorielle Validität bis auf die Strategie *Wiederholung* nachgewiesen werden. Die bisherigen Analysen zum LIST weisen jedoch einen entscheidenden Nachteil auf. Als Methode der Faktorbestimmung wurde sowohl von Wild und Schiefele (1994) als auch von Boerner et al. (2005) die Hauptkomponentenanalyse verwandt. Diese entspricht nicht dem *Common-factor*-Modell von Thurstone (1947) und spiegelt keine der Itembeantwortung zugrunde liegende latente Dimension wider (s. Kapitel 3).

In der vorliegenden Arbeit wird der LIST in der ursprünglichen Version von Wild und Schiefele (1994) mit 77 Items verwendet. Wegen des angeführten Nachteils der Verwendung von Hauptkomponentenanalysen zur Skalenkonstruktion werden hier die LIST Skalen nicht ungeprüft übernommen. Um die faktorielle Validität und die Gültigkeit des Messmodells der einzelnen LIST Skalen sicherzustellen, werden diese vor ihrer Verwendung mithilfe einer konfirmatorischen bzw. Hauptachsenfaktorenanalyse auf faktorielle Validität untersucht.

1.7 Epistemologische Überzeugungen als latente Variablen

Bevor in diesem Kapitel methodologische und wissenschaftstheoretische Aspekte vorgestellt werden, die für die vorliegende Arbeit relevant sind, werden in einem kurzen Resümee die bisher angesprochenen Sachverhalte zusammengefasst. In den vorangehenden Kapiteln wurden vier grundlegende Ansätze bei der Erforschung epistemologischer Überzeugungen vorgestellt. Insbesondere wurde dabei auf Modelle aus dem persönlichkeitspsychologischen Ansatz eingegangen. Für die beabsichtigten Untersuchungen sind dabei insbesondere die Modelle von Schommer (1990) und Moschner sowie Gruber (2005) relevant. Im Anschluss hieran wurden verschiedene Ansätze epistemologischer Überzeugungen dargestellt, die für spätere Überlegungen wichtig sein werden. Die pädagogische Relevanz epistemologischer Überzeugungen wurde abschließend anhand ihrer Beziehung zu Lernerfolg und Lernstrategien dargestellt.

Da insbesondere die persönlichkeitspsychologischen Modelle im Folgenden relevant sind, werden an dieser Stelle die methodologischen Grundlagen und wissenschaftstheoretischen Aspekte näher ausgeführt, die dem persönlichkeitspsychologischen Ansatz zugrunde liegen. Im Vordergrund stehen dabei jedoch nicht die statistischen Verfahren, die bei den empirischen Untersuchungen angewendet und ausführlich in Kapitel 3 abgehandelt werden. Insbesondere wird erarbeitet, welche Aussagen aus dem persönlichkeitspsychologischen Ansatz gewonnen werden können und was unter dem Begriff *latente Variable* verstanden wird. Anschließend wird aufgezeigt, welche Folgerungen sich aus diesem Verständnis für die Forschung ergeben.

Der hier vertretene Ansatz epistemologischer Überzeugungen konzentriert sich auf die pädagogisch-psychologische bzw. persönlichkeitspsychologische Perspektive, die zuvor mit den Ansätzen von Schommer (1990), Hofer und Pintrich (1997) sowie dem Ansatz von Moschner

und Gruber (2005) beschrieben wurde. Die Konzentration auf den persönlichkeitspsychologischen Ansatz beruht hauptsächlich auf drei pragmatischen Überlegungen:

- Der persönlichkeitspsychologische Ansatz ist der am weitesten verbreitete Ansatz. Es liegt eine Vielzahl von Fragebögen vor, die zur Erfassung interindividueller Unterschiede herangezogen werden. Eines der am häufigsten eingesetzten Verfahren ist der Fragebogen von Schommer (1990). Allerdings sind aus methodischer Sicht keine adäquat durchgeführte Evaluation von Schommers Modell bzw. Studien zur psychometrischen Qualität von Schommers Fragebogen bekannt. Zudem ist keine offiziell publizierte deutsche Version von Schommers EBQ verfügbar. Mit dem Modell von Moschner und Gruber ist ein umfassendes Modell epistemologischer Überzeugungen aufgestellt und ein dazugehöriger Fragebogen entwickelt worden. Zu diesem Modell und dem dazugehörigen Fragebogen fehlt allerdings eine empirische Überprüfung und es bedarf einer Evaluation der empirischen Resultate.
- Er bietet mit seiner Verwendung von Items und Ratingskalen eine Möglichkeit der ökonomischen Diagnostik epistemologischer Überzeugungen. Daher können mithilfe des persönlichkeitspsychologischen Ansatzes Vergleiche in der Ausprägung epistemologischer Überzeugungen zwischen verschiedenen Personen getroffen werden und es besteht die Möglichkeit, Zusammenhänge mit anderen Variablen zu untersuchen. Diese Möglichkeit besteht zwar ebenfalls bei den entwicklungspsychologischen Ansätzen mit den dort vorherrschenden Interviewverfahren, der Einsatz von Skalen gestaltet sich aber weitaus einfacher.
- Andere Ansätze, wie z. B. der von Hammer und Elby (2002), sind im Hinblick auf die Methodik weniger ausgearbeitet. Oder es wird, wie z. B. in dem Ansatz von Royce, zwar ein alternatives kognitives Modell vorgestellt, die Erfassung epistemologischer Überzeugungen erfolgt jedoch mithilfe von Ratingskalen, sodass in methodischer Hinsicht ein persönlichkeitspsychologisches Vorgehen vorliegt.

Charakteristisch für die hier bevorzugten persönlichkeitspsychologischen Ansätze ist die Erfassung von epistemologischen Überzeugungen durch Selbstbeurteilungen anhand von Ratingskalen, die von vielen Individuen bearbeitet werden. Im Anschluss werden die erfassten Daten mithilfe von Faktoren- und Korrelationsanalysen ausgewertet und es werden die Beziehungen der verschiedenen Dimensionen epistemologischer Überzeugungen untereinander bzw. die Beziehungen zu anderen Konstrukten (z. B. Lernstrategien) bestimmt. Dieses Vorgehen der Erfassung von Persönlichkeitseigenschaften wird als R-Technik bezeichnet

(Loehlin, 1987). Gleichzeitig wird damit allerdings auch der Schwerpunkt auf die Bestimmung interindividueller Differenzen gelenkt und damit von einer Konsistenz epistemologischer Überzeugungen im Sinn des Trait-Paradigmas ausgegangen (Asendorpf, 1995, 1999). Die Konsistenz folgt aus der Verwendung der R-Technik, bei der über die Messsituationen hinweg generalisiert wird, sodass diese nicht in die Betrachtung eingehen. Überdies wird die Betrachtung auf Ebene der Population verschoben wird, d. h. die mittels korrelativer Techniken ermittelten Parameter sind immer Parameter auf Populationsebene (Asendorpf, 1995). Wobei der Begriff Population ist nicht mit dem Begriff aus der Statistik gleichzusetzen, Population meint hier lediglich eine Menge von Individuen und dass situative Aspekte ausgeblendet werden. Die Betrachtung erfolgt somit im Hinblick auf die interindividuellen Unterschiede der Populationsmitglieder. Nach Stern (1911) handelt es sich daher um die Perspektive der Differenziellen Psychologie im engeren Sinn.²⁶ Stern unterscheidet die Variationsforschung und die Korrelationsforschung. Die Variationsforschung untersucht die Variation eines Merkmals an mehreren Individuen und die Korrelationsforschung untersucht die Korrelation mehrerer Merkmale bei verschiedenen Individuen. Der Zusammenhang epistemologischer Überzeugungen und anderer Konstrukte gehört nach Stern in die Korrelationsforschung.

Die Folge der Konzentration auf interindividuelle Differenzen besteht darin, dass die intraindividuelle Variation zwischen Situationen sowie eine eventuell vorhandene Interaktion zwischen Situation und Person nicht berücksichtigt werden, aus methodischer Sicht werden diese dem Messfehler zugerechnet. Die gefundenen interindividuellen Zusammenhänge sind somit in der Regel nicht auf intraindividuelle Zusammenhänge zu übertragen. Dies wäre nur bei funktioneller Homogenität der Fall (Valsiner, 1986). Dabei entsprechen die interindividuellen Zusammenhänge auf Ebene der Population den intraindividuellen Zusammenhängen auf Ebene der Person. Funktionelle Homogenität muss jedoch durch geeignete Verfahren nachgewiesen werden (Hamaker, Dolan & Molenaar, 2005).

Neuerdings werden in Bezug auf die funktionelle Homogenität drei unterschiedliche Arten von Konstrukten unterschieden (Ellis & van den Wollenberg, 1993):

- Zum einen gibt es *lokal homogene* Konstrukte, bei denen sich die auf der interindividuellen Ebene ermittelten Zusammenhänge auch auf der intraindividuellen Ebene wiederfinden. Nach Borsboom, Mellenbergh und van Heerden (2003) liegen solche Zusammenhän-

²⁶ Stern (1911) definiert in seinem Werk die Differenzielle Psychologie als Zweig der Psychologie und unterscheidet weiter nach universellen, speziellen und differenziellen Fragestellungen. Die differenzielle Fragestellung bildet die Differenzielle Psychologie im engeren Sinn (Asendorpf, 1995).

ge nur dann vor, wenn wiederholte Beobachtungen denselben vorgenommen werden und die daraus abgeleitete Zusammenhgangsstruktur die Gleiche ist, wie die Zusammenhgangsstruktur einer Messung mithilfe der R-Technik ergeben würde (Borsboom et al., 2003). Die wiederholte Erfassung einer oder mehrerer Variablen an derselben Person entspricht der P-Technik (Loehlin, 1987). Im Falle lokal homogener Konstrukte ist es erlaubt, Hypothesen über intraindividuelle Zusammenhänge durch die Analyse von interindividuellen Zusammenhängen zu prüfen.

- Zum anderen gibt es *lokal heterogene* Konstrukte, bei welchen sich eine interindividuelle Zusammenhangstruktur auf der Stichprobenebene finden lässt, aber es existieren unterschiedliche intraindividuelle Zusammenhangstrukturen.
- Darüber hinaus gibt es *lokal irrelevante* Konstrukte, bei denen nicht entschieden werden kann, ob es sich um lokal homogene oder lokal heterogene Konstrukte handelt. Nach Borsboom et al. (2003) ist dies insbesondere bei Konstrukten der Fall, die als stabile Traits konzipiert sind. Das charakteristische Merkmal dieser Traits besteht darin, dass sich diese bei wiederholten Messungen nicht verändern. Dadurch wird bei wiederholten Messungen auch keine Varianz auf der intraindividuellen Ebene erzeugt, sodass auch keine intraindividuelle Zusammenhangstruktur gefunden werden kann. Lediglich auf der interindividuellen Ebene wird durch die Messung an mehreren Individuen Varianz erzeugt, welche es erlaubt, die Zusammenhangstruktur mithilfe der R-Technik zu erkennen.

Durch die Verwendung der R-Technik, die keine Messwiederholung beinhaltet und folglich von einer Konsistenz bzw. einer Traitkonzeption epistemologischer Überzeugungen ausgeht, ist es in diesem Ansatz daher nur möglich, Aussagen über interindividuelle Differenzen zu tätigen. Intraindividuelle Zusammenhänge und Prozesse sind durch die angewandten Methoden nicht zu untersuchen, insbesondere kann durch die Anwendung der R-Technik die Frage nach der funktionellen Homogenität nicht entschieden werden.

Eine weitere Folge der Methode zur Erfassung epistemologischer Überzeugungen mit Selbstbeurteilung auf Ratingskalen besteht darin, dass diese nicht direkt erfasst werden können. In diesem Zusammenhang wird zwischen manifesten und latenten Variablen unterschieden. Manifeste Variablen sind beobachtbar und die Kenntnis ihrer Ausprägungen erlaubt einen Rückschluss auf die Ausprägung der latenten Variablen, die selbst nicht beobachtbar sind. Es handelt sich damit um eine indirekte Messung (Meyer, 2004). Dabei wird davon ausgegangen, dass die latente Variable kausal das Antwortverhalten der betrachteten Personen bei den ma-

nifesten Variablen beeinflusst (Borsboom, Mellenberg & van Heerden, 2003, 2004, Borsboom, Cramer, Kievit, Zand Scholten & Franic, 2009a). Nach Edwards und Bagozzi (2000) handelt es sich somit um ein direkt reflexives Messmodell. Die Verbindung von nicht direkt beobachtbaren latenten Variablen und den durch sie bedingten manifesten Variablen wird durch ein solches Messmodell hergestellt. Ein Messmodell ist eine Funktion, welche die latente Variable mit den manifesten Variablen in Beziehung setzt (Borsboom, 2008). Wie in Kapitel 3 bei der Beschreibung der statistischen Methoden aufgezeigt wird, handelt es sich bei dieser Funktion um das *Common-factor*-Modell von Thurstone (1947), bei dem die manifeste Variable eine lineare Funktion der latenten Variablen und einer Fehlervariablen ist. Die latente Variable wird dabei ebenso wie die manifesten Variablen als kontinuierlich aufgefasst. Somit handelt es sich um die Art latenter Variablen, wie sie üblicherweise in einer faktorenanalytischen Betrachtung zugrunde liegen. Aus mathematischer Sicht bestimmen die latenten Variablen die Kovarianzstruktur der manifesten Variablen (Loehlin, 1987). Die sich aus der Annahme latenter Variablen ergebenden Implikationen werden im Folgenden in Anlehnung an die Darstellung von Borsboom et al. (2003) dargestellt.

Formal stellt eine latente Variable eine mathematische Funktion dar (Borsboom, 2008), die das reflexive Messmodell repräsentiert. Diese mathematische Darstellung zeigt die Syntax, an die sich eine semantische Sichtweise anschließt, welche die Bedeutung des Formalismus angibt. In diesem Fall ist dies die *Repeated-subjects-sampling*-Interpretation (Borsboom et al., 2003, Meredith, 1993), die der zuvor vorgestellten R-Technik entspricht. Zu der formalen Sichtweise kommt ein empirischer Aspekt hinzu, der das mathematische Modell mit den erhobenen Daten in Beziehung setzt. Hierzu wird geprüft, ob die durch das Modell implizierte Kovarianzstruktur mit der Kovarianzstruktur der empirischen Daten übereinstimmt. Ergibt sich eine Passung, wird in der Regel das formale Modell als richtig angesehen. Allerdings kann aufgrund dieser Passung keine Aussage über die Natur der latenten Variablen gemacht werden. Im Allgemeinen gibt es viele formale Modelle, welche zu den empirischen Daten passen, diese werden als äquivalente Modelle bezeichnet (Stelzl, 1986). Äquivalente Modelle weisen die gleiche Kovarianzstruktur auf und zeigen ebenfalls die gleiche Modellpassung und sind daher mittels statistischer Methoden nicht zu unterscheiden (MacCallum, Wegener, Uchino & Fabrigar, 1993). Daraus folgt, dass das formale Modell keine Aussagen über die Eigenschaft der latenten Variablen erlaubt.

Benötigt wird somit also noch eine Ontologie der latenten Variablen, damit Aussagen über die Zusammenhänge von latenten und manifesten Variablen getroffen werden können. Insbesondere muss dabei die Auffassung berücksichtigt werden, dass die latente Variable die Kovarianzstruktur der manifesten Variablen kausal bedingt. Um von Kausalität ausgehen zu können, ist es nach Shadish, Cook und Campbel (2002) erforderlich, dass die Variation der latenten Variablen der Variation der manifesten Variablen vorausgeht. Nach Borsboom et al. (2003) erlaubt die realistische Position nach Devitt (1991), die erforderlichen ontologischen Aussagen über die latente Variable zu treffen. Borsboom et al. (2003) bezeichnen diese realistische Position als Entitätenrealismus.

Der Entitätenrealismus ist eine Ausprägung des Realismus, der sich durch drei Annahmen kennzeichnen lässt:

1. Es gibt Theorien und diese sind entweder wahr oder falsch.
2. Es gibt Entitäten, die sich mithilfe von Theorien beschreiben lassen.
3. Diese Entitäten sind kausal für das Zustandekommen von empirisch beobachtbaren Phänomenen verantwortlich.

Es muss nun gezeigt werden, wie die Grundannahmen des Entitätenrealismus es gestatten, die Ontologie latenter Variablen zu bestimmen. Hierzu sind die zweite und die dritte Annahme relevant. Zuvor wurde festgehalten, dass es sich bei dem benutzten Messmodell, welches auf dem *Common-factor*-Modell von Thurstone (1947) beruht, um ein direkt reflexives Messmodell handelt. Das direkt reflexive Messmodell geht davon aus, dass eine Veränderung der latenten Variable zu Veränderungen in den manifesten Variablen führt. Veränderungen der manifesten Variablen setzen also voraus, dass sich die latente Variable verändert hat. In den üblichen mathematischen Modellen stellen manifeste Variablen Regressionen auf die latente Variable dar. Wird nun die latente Variable auf einer Ausprägungsstufe konstant gehalten, so führt dies dazu, dass auch die manifesten Variablen keine Varianz mehr aufzeigen. Diese Eigenschaft wird als lokale Unabhängigkeit bezeichnet (vgl. Bühner, 2006). Da es sich bei latenten Variablen um nicht direkt beobachtbare Variablen handelt, die erst durch eine indirekte Messung über die manifesten Variablen erschlossen werden, kann keine unmittelbare Konstanthaltung der latenten Variablen erfolgen. Die kausale Beeinflussung der manifesten Variablen durch die latente Variable kann allerdings kontrafaktisch gezeigt werden. Bestünde die Möglichkeit, die latente Variable konstant zu halten, so gäbe es zwischen den manifesten Variablen keine Kovarianz. Lokale Unabhängigkeit gestattet es, die Variation der latenten Variablen als Ursache der Variation der manifesten Variablen anzusehen (Pearl, 2000). Der

Nachweis der Kausalität durch den kontrafaktischen Schluss entspricht der dritten Annahme. Mit der Annahme einer Kausalbeziehung geht aber auch die Annahme von Entitäten einher. Wird zwischen der latenten Variable und den manifesten Variablen eine Kausalbeziehung angenommen, so wird damit auch eine regelmäßige, gesetzmäßige Verknüpfung unterstellt. Dies ist sowohl bei materiellen als auch nicht-materiellen Prozessen möglich (Bunge, 1987, Westermann 2000). Latente Variablen in diesem Sinn sind demnach nicht materielle Entitäten, auf die durch die Regelmäßigkeiten in den manifesten Variablen geschlossen werden kann. Borsboom et al. (2003) verdeutlichen dies an einem Beispiel. Dominante Menschen werden Items zur Erfassung von dominantem Verhalten zustimmen und nicht dominante Menschen werden diesen Items nicht zustimmen. Damit einher geht aber die Annahme der Eigenschaft *Dominanz*, welche das Antwortverhalten verursacht. *Dominanz* ist in diesem Zusammenhang die gesuchte Entität. Kontrafaktisch kann nun argumentiert werden: Gäbe es die Eigenschaft *Dominanz* nicht, so würde sich keine systematische Variation der manifesten Variablen ergeben. Ebenso würde die Variation der manifesten Variablen ausbleiben, wenn die Eigenschaft *Dominanz* konstant gehalten würde. Dies entspricht somit der zweiten Annahme des Entitätenrealismus. Latente Variablen können damit als nicht materielle Eigenschaften angesehen werden, die überdies durch ihre Wirkung auf manifeste Variablen messbar sind.

Die dargelegte Methodologie des persönlichkeitspsychologischen Ansatzes epistemologischer Überzeugungen führt zu einigen Folgerungen, die bei der Interpretation von Forschungsergebnissen berücksichtigt werden müssen: Einerseits können nicht ohne Weiteres Aussagen über die intraindividuelle Struktur von epistemologischen Überzeugungen getroffen werden können. Dies wäre nur im Fall lokal homogener Konstrukte möglich und ergibt sich überdies aus der Verwendung der R-Technik bzw. der *Repeated-subjects-sampling*-Interpretation, die sich auf Zusammenhänge auf Populationsebene bezieht (vgl. Borsboom et al, 2003). Andererseits ergibt sich daraus, dass korrelative Zusammenhänge mit anderen Konstrukten, wie sie z. B. im Rahmen einer Validierung bestimmt werden, sich ebenfalls nur auf interindividuelle Zusammenhänge beziehen. Auf der Grundlage der Zusammenhänge mehrerer Konstrukte kann nicht auf intraindividuelle Zusammenhänge dieser Konstrukte geschlossen werden. An einem Beispiel erläutert bedeutet dies, dass z. B. ein positiver Zusammenhang von epistemologischen Überzeugungen und Textverständnis (vgl. Schommer, 1993a) nur auf Populationsebene, nicht aber auf individueller Ebene gilt. D. h. es kann nicht geschlossen werden, dass dieser Zusammenhang bei wiederholter Messung von epistemologischen Überzeugungen und Textverständnis bei einer Person zu finden ist (vgl. Asendorpf, 1995).

Auch ergeben sich weiterhin auch Konsequenzen für den Fall, dass epistemologische Überzeugungen zu diagnostischen Zwecken erfasst werden sollen. Die Schwerpunktsetzung auf interindividuelle Unterschiede impliziert, dass mit Messinstrumenten, welche auf der R-Technik basieren, nur Vergleiche zwischen Personen vorgenommen werden können. Die Messung epistemologischer Überzeugungen ermöglicht somit nur den Vergleich der Ausprägungsgrade von epistemologischen Überzeugungen zwischen verschiedenen Individuen. Daraus ergibt sich auch eine normorientierte Messung, bei der das Individuum mit einem Kennwert der jeweiligen Referenzpopulation verglichen wird. Es wird also der Status der Ausprägung epistemologischer Überzeugungen zu einem Zeitpunkt bei einer Person festgehalten. Da im Trait-Paradigma von einer gewissen Konstanz epistemologischer Überzeugungen ausgegangen wird, liegt also Statusdiagnostik vor (Amelang & Zielinski, 2002). Diese Statuserfassung kann diagnostischen oder prognostischen Zwecken dienen.

Zum Abschluss dieses Kapitels soll die bisher dargestellte erkenntnistheoretische Sichtweise hinsichtlich latenter Variablen noch kritisch reflektiert werden. Die hier dargestellte Position, die sich weitestgehend auf die Arbeit von Borsboom et al. (2003) stützt, sieht latente Variablen unter einer realistischen Perspektive als real existierende Entitäten an, welche die Korrelationen von manifesten Variablen bedingen. Gegen diese Position ist allerdings auch Kritik vorgetragen worden. Lee (2012) widerspricht der Auffassung, dass latente Variablen als Ursache für die Korrelationen von manifesten Variablen angesehen werden können. Lees Argumentation baut auf zwei Argumenten auf:

1. Wissenschaftler verhalten sich nicht so, als wäre die Variation der latenten Variablen die Ursache für die Variation der manifesten Variablen, wenn sie z. B. semantisch gleiche Items hinzufügen, um die Reliabilität einer Skala zu erhöhen.
2. Es ist nicht möglich, die latenten Variablen unabhängig von den manifesten Variablen zu erfassen und somit auch nicht möglich, Ursachen von Wirkungen zu trennen.

Das erste Argument spricht nicht gegen eine realistische Sichtweise latenter Variablen. Zum einen kann das Verhalten von Wissenschaftlern nicht als Argument gegen einen bestimmten erkenntnistheoretischen Standpunkt dienen. Es ist weder zwingend erforderlich, dass sich in bestimmten Verhaltensweisen eine widersprechende erkenntnistheoretische Sichtweise widerspiegelt, noch dass die betreffenden Wissenschaftler überhaupt irgendeine erkenntnistheoretische Sichtweise vertreten. Zum anderen kann auch eine – wenig reflektierte – methodische Praxis ebenfalls nicht als Argument gegen eine erkenntnistheoretische Position vorgebracht werden, nur der Grund bzw. die Gründe für diese Praxis können als Argument dienen.

Das zweite Argument von Lee (2012) wiegt dagegen stärker, spricht aber letztendlich ebenfalls nicht gegen die vorgestellte Sichtweise latenter Variablen. Das Argument, dass zwei Variablen unabhängig voneinander definiert sein müssen, wird von Trafimow und Rice (2009) untersucht.²⁷ Die Autoren benutzten dort das zweite Newton'sche Gesetz, um zu zeigen, dass die unabhängige Definition zweier Begriffe nicht notwendig bzw. erforderlich ist. Das zweite Gesetz sagt aus, dass Kraft gleich dem Produkt aus Masse und Beschleunigung ist. Danach lässt sich einer der Begriffe Kraft, Masse oder Beschleunigung jeweils in Abhängigkeit von den anderen Begriffen definieren. Dies führt jedoch dazu, dass die klassische Mechanik zu einer zirkulären Theorie wird. Um dies zu umgehen müsste mindestens einer der Begriffe Kraft, Masse oder Beschleunigung unabhängig von den anderen Begriffen definiert werden. Nach Trafimow und Rice (2009) stellt eine solche Zirkularität kein Hindernis dar. Würde nämlich, wie gefordert, ein Begriff unabhängig von den anderen definiert werden, so würde dies bedeuten, dass in diese Definition wenigstens ein theoretischer Term eingeführt werden würde. Dieser theoretische Term wäre jedoch wieder durch einen weiteren theoretischen Term zu definieren. Und dieser neue theoretische Term wäre ebenfalls wieder durch einen theoretischen Term zu definieren, so dass eine infinite Kette von zu definierenden theoretischen Termen folgen würde. Trafimow und Rice (2009) schlussfolgern daraus, dass es ausreichend ist, wenn eine Definition Begriffe enthält, die nicht weiter zu definieren sind und damit sogenannte undefinierte Grundbegriffe enthält (Westermann, 2000).

Aus dem Beispiel des zweiten Newton'schen Gesetzes kann überdies noch eine Analogie abgeleitet werden, welche die eigene Argumentation einer realistischen Sichtweise latenter Variablen stützt. Betrachtet man die Erdanziehungskraft, so ergibt sich aus dem zweiten Gesetz, dass die Erdanziehungskraft gleich dem Produkt der Masse eines Objekts mal der Schwerebeschleunigung der Erde ist, wobei die Schwerebeschleunigung diejenige Beschleunigung darstellt, die ein Objekt durch die Erdanziehungskraft erfährt und die überall auf der Erdoberfläche (annähernd) gleich ist. Die Erdanziehungskraft kann nun als latente (nicht direkt beobachtbare) Variable gedacht werden, die ursächlich dafür ist, dass ein Objekt auf der Erde eine Masse hat. Die Masse eines Objekts wird somit durch die Erdanziehungskraft bedingt. Aus kontrafaktischer Sicht würde sich daraus Folgendes ergeben: Ohne die Erdanziehung hat

²⁷ In dem Artikel von Trafimow und Rice (2009) werden verschiedene Theorien aus den letzten Jahrhunderten kurz vorgestellt, anschließend stellen die Autoren kurze, fiktive Gutachten dar, die begründen, warum die jeweilige Theorie nicht für eine Veröffentlichung in einer Zeitschrift in Betracht kommt. Beispiele für Theorien, welche von den Autoren als Beispiel herangezogen werden, sind unter anderem das heliozentrische System von Galilei, die Mechanik Newtons, die Darstellung des Blutkreislaufs von William Harvey und die Spezielle Relativitätstheorie von Einstein.

ein Objekt keine Masse.²⁸ Darüber hinaus wäre es in diesem Beispiel möglich die Wirkung der Erdanziehung manipulativ zu „verändern“. Natürlich ist es nicht möglich die Erdanziehung zu verändern, allerdings haben andere Himmelskörper wie beispielsweise der Mond eine andere Schwerebeschleunigung als die Erde. Zur Veränderung der Schwerebeschleunigung müsste ein Gegenstand daher auf den Mond gebracht werden. Auf dem Mond beträgt die Schwerebeschleunigung nur ein Sechstel der Schwerebeschleunigung auf der Erde, somit hat ein Objekt dort nur ein Sechstel der Masse, die es auf der Erde hätte. Zusammenfassend lässt sich damit festhalten, dass trotz der Einwände von Lee (2012) die in diesem Kapitel dargestellte Auffassung latenter Variablen und den sich daraus ergebenden Konsequenzen beibehalten werden.

²⁸ Der Unterschied zu dem weiter oben angeführten Beispiel mit der Eigenschaft Dominanz ist, dass dort eine kontrafaktische Argumentation dazu führt, da die Variation der manifesten Variablen an eine Menge von Personen nicht mehr gegeben ist, während in diesem Beispiel die Masse von einem Objekt nicht mehr vorhanden ist; es geht in diesem Beispiel somit nicht um die Variation der Masse.

2. Forschungsfragen

Bevor die konkreten Forschungsfragen betrachtet werden, soll die übergeordnete Zielsetzung aufgezeigt werden. Da es eine Vielzahl von Verfahren zur Erfassung epistemologischer Überzeugungen im deutschen Sprachraum gibt (Priemer, 2006), von denen die meisten jedoch methodisch nicht stringent entwickelt wurden, ist das Hauptziel die *methodisch exakte Untersuchung deutschsprachiger Instrumente zur Erfassung epistemologischer Überzeugungen*. Der Schwerpunkt liegt auf dem EBQ-Fragebogen von Schommer (1990) in einer deutschen Version sowie auf den EPI-Fragebogen von Moschner und Gruber. Die Gründe für die Auswahl dieser beiden Verfahren sind beim EBQ seine Bekanntheit und die Häufigkeit seiner Anwendung in empirischen Untersuchungen. Der EPI-Fragebogen wird wegen seiner theoretischen Fundierung und Neuheit betrachtet.

Speziell analysiert werden zwei Bereiche, der erste Bereich deckt Untersuchungen zur faktoriellen Struktur und den psychometrischen Eigenschaften der daraus abgeleiteten Skalen des EBQ von Schommer (1990) ab. Der zweite Bereich beschäftigt sich mit der faktoriellen Struktur, den psychometrischen Eigenschaften der abgeleiteten Skalen sowie der Validität des EPI-Fragebogens von Moschner und Gruber. Die jeweiligen Fragestellungen werden im Folgenden näher vorgestellt und zusammenfassend dargestellt.

Bei dem EBQ soll in einem ersten Schritt die Modellvorstellung von Schommer geprüft werden, dessen festgesetzte Zuordnung von Item zu Skalen empirisch geprüft wird. Diese Fragestellung wird als essenziell betrachtet, da bei der Anwendung dieses Fragebogens von Schommer selbst die Verwendung dieser festgelegten Skalen empfohlen wird. Dieses Vorgehen wurde auch in vielen Publikationen umgesetzt. Umso wichtiger ist es, das Modell und die Praxis der Skalenbildung empirisch zu überprüfen, vor allem vor dem Hintergrund der Kritik an dem EBQ, die in der im vorhergehenden Kapitel referierten Literatur vorgebracht wurde.

Die zweite Fragestellung beschäftigt sich mit der faktoriellen Struktur des EBQ. Das damit verfolgte Ziel ist die Ermittlung der Faktorenstruktur mittels einer Faktorenanalyse auf Itemebene. Die Klärung dieser Struktur ist insofern relevant, als dass bis zu diesem Zeitpunkt nur einer Studie bekannt ist, in der eine faktorenanalytische Untersuchung der ursprünglichen 63 Items des EBQ vorgenommen wurde (vgl. die tabellarische Zusammenstellung der relevanten Arbeiten zu dieser Thematik in Buehl, 2008). Entweder wurde in den relevanten Arbeiten eine Faktorisierung mit den Subsets von Schommer vorgenommen oder es wurden die Items

verändert oder die Arbeiten weisen im Hinblick auf die verwendete faktorenanalytische Methodik Mängel auf. Die Prüfung der faktoriellen Struktur des Schommer-Fragebogens gliedert sich dabei in mehrere Teilschritte. Zunächst muss geklärt werden, wie viele Faktoren der EBQ beinhaltet. Dazu müssen Methoden zum Einsatz kommen, welche es gestatten, die Anzahl der zugrunde liegenden Faktoren möglichst objektiv zu bestimmen. Ist die Zahl der Faktoren bestimmt, steht die Frage nach der faktoriellen Struktur des Modells im Raum. Dazu gehören die Ladungen der Items auf den jeweiligen Faktoren sowie die Beziehungen der Faktoren untereinander. Von Interesse ist hier, inwiefern die Faktoren miteinander korreliert sind. Damit soll die Formulierung „mehr oder weniger“ voneinander unabhängiger Dimensionen epistemologischer Überzeugungen von Schommer (1990) auf eine empirische Basis gestellt werden. Diese Fragestellung ist in erster Linie explorativ und wird daher mithilfe von explorativen Faktorenanalysen bearbeitet.

Die nächsten Fragen in Bezug auf den Schommer-Fragebogen, die beantwortet werden sollen, zielen auf die psychometrischen Eigenschaften der aus den Faktoren abgeleiteten Skalen ab. Die Frage nach den psychometrischen Eigenschaften der gebildeten Skalen beschränkt sich nicht nur auf die Bestimmung der Reliabilität der Skalen, sondern schließt auch die Frage nach der Äquivalenz der Items mit ein. D. h. es wird geprüft, ob τ -äquivalente oder parallele Items zur Erfassung der jeweiligen Skala vorliegen.

Zusammenfassend lassen sich für den Fragebogen von Schommer folgende Fragestellungen festhalten:

- *Lässt sich das Modell von Schommer durch eine konfirmatorische Faktorenanalyse empirisch bestätigen?*
- *Wie ist die faktorielle Struktur des Schommer-Fragebogens bei Faktorisierung auf der Ebene der einzelnen Items?*
- *Welche Skalen lassen sich aus den Ergebnissen der Faktorenanalyse ableiten und welche psychometrischen Eigenschaften besitzen diese Skalen?*

Die Forschungsfragen zum EPI-Fragebogen betreffen ebenfalls zuerst die faktorielle Struktur. Insbesondere muss geklärt werden, ob die von Moschner und Gruber gefundene Faktorenstruktur bestätigt werden kann. Diese Frage ist in erster Linie konfirmatorischer Natur und aus diesem Grund werden die Skalen des EPI-Fragebogens mithilfe konfirmatorischer Faktorenanalysen auf ihre Modellpassung untersucht, wobei die Passung des Gesamtmodells und der

einzelnen Faktoren gesondert betrachtet werden soll. Von besonderem Interesse dabei ist, ob die hier vorliegenden Daten des EPI-Fragebogens zu gleichen oder ähnlichen Ergebnissen führen, die von Moschner und Gruber (im Druck) bei der Weiterentwicklung des EPI-Fragebogens zu dem FEE-Fragebogen vorgelegt wurden. Sollte eine Modifikation der Faktoren von Moschner und Gruber (2005) notwendig werden, müssen diese kreuzvalidiert werden. Daran anschließend werden die psychometrischen Eigenschaften der gebildeten Skalen untersucht. Auch hier stehen die Frage nach der Äquivalenz der Items sowie die Bestimmung der Reliabilität im Mittelpunkt. Analog dem Vorgehen bei der Untersuchung der Skalen des EBQ wird untersucht, ob τ -äquivalente oder parallele Items vorliegen und wie hoch die Reliabilitäten der Skalen sind.

Weiterhin wird für den EPI-Fragebogen eine Konstruktvalidierung durchgeführt, die in diesem Fall die konkurrente Validität in Bezug auf Lernstrategien umfasst. Dabei wird gemäß Darstellung der Forschungsergebnisse in Kapitel 1.6 davon ausgegangen, dass zwischen epistemologischen Überzeugungen und Lernstrategien ein Zusammenhang besteht. Allerdings ist es an dieser Stelle nicht möglich, die Zusammenhangshypothesen der EPI-Faktoren mit Lernstrategien a priori festzulegen, da diese Hypothesen von den vorausgegangenen Prüfungen der Faktoren des EPI-Fragebogens abhängen. Daher ist hier nur eine allgemeine Zusammenhangshypothese aufzustellen, nach der es einen Zusammenhang zwischen epistemologischen Überzeugungen und Lernstrategien gibt.

Zusammenfassend lassen sich für den EPI-Fragebogen folgende Fragestellungen festhalten:

- *Lässt sich die Faktorenstruktur des EPI-Fragebogens mithilfe einer konfirmatorischen Faktorenanalyse bestätigen?*
- *Lassen sich die einzelnen Skalen des EPI-Fragebogens mithilfe einer konfirmatorischen Faktorenanalyse bestätigen und welche Ähnlichkeiten und Unterschiede gibt es zwischen diesen Ergebnissen und dem FEE von Moschner und Gruber (im Druck)?*
- *Lassen sich ggf. notwendige Modifikationen der Skalen des EPI mithilfe einer Kreuzvalidierung absichern?*
- *Welche psychometrischen Eigenschaften besitzen die Skalen des EPI-Fragebogens?*
- *Lassen sich zwischen den EPI-Faktoren und Lernstrategien positive Zusammenhänge nachweisen?*

3. Methoden

Dieser Teil ist den Methoden gewidmet, die für die Bearbeitung der formulierten Forschungsfragen benötigt werden. Dabei handelt es sich in erster Linie um Methoden aus dem Bereich der Faktorenanalyse und der Psychometrie. Faktorenanalytische Methoden werden zur Analyse der Struktur der Fragebögen benötigt, wohingegen psychometrische Methoden benötigt werden, um die Eigenschaften der Skalen zu ermitteln, die mithilfe der Faktorenanalyse abgeleitet wurden.

Die Darstellung der angewandten Methoden beginnt bei der explorativen und konfirmatorischen Faktorenanalyse, wobei sich die Darstellung der explorativen Faktorenanalyse an Klopp (2010) orientiert. Hieran schließen eine Darstellung des Begriffs der Validität sowie eine Skizze der Regressionsanalyse als lineares Strukturgleichungsmodell an, die zur Untersuchung der konkurrenten Validität der auf faktorenanalytischem Weg ermittelten Skalen verwendet wird. Die Strategie des Methodeneinsatzes für die Beantwortung der Forschungsfragen sowie die Konsequenzen, die sich aus den benutzten Methoden ergeben, werden jeweils nach der Besprechung einer Methode in einem separaten Kapitel diskutiert.

3.1 Explorative Faktorenanalyse

Die explorative Faktorenanalyse ist ein struktorentdeckendes Verfahren, das üblicherweise dann zum Einsatz kommt, wenn die Konstrukte, welche die Antwort auf verschiedene Items beeinflussen, noch unbekannt bzw. wenig ausgearbeitet sind. Die Grundannahme der Faktorenanalyse lautet: Der Wert, welcher eine Person auf einem Item ausweist, setzt sich additiv aus einer gewichteten Summe der Faktoren bzw. latenten Variablen zusammen, die Antworten der Person auf diesem Item werden durch die Faktoren bzw. latenten Variablen bedingt. Die Grundannahme der Faktorenanalyse entspricht dem Modell gemeinsamer Faktoren (*common factor model*) nach Thurstone (1947). Dabei wird die Varianz der manifesten Werte aufgeteilt in die Varianz, welche durch die latenten Faktoren verursacht wird, und in eine Residualvarianz. Die Residualvarianz ergibt sich aus der Varianz der beobachteten Variablen, die nicht durch die gemeinsamen Faktoren verursacht wird, und aus unsystematischer Fehlervarianz, die weder auf die Faktoren noch auf das Item zurückzuführen ist (Brown, 2006). Das Besondere an der explorativen Faktorenanalyse bzw. dem Modell gemeinsamer Faktoren ist, dass jede der angenommenen latenten Variablen einen Einfluss auf die manifesten Variablen hat.

Aus den beobachteten Werten müssen anschließend auf einem analytischen Weg die Faktorladungen sowie die Faktorwerte ermittelt werden. In der Regel sind dabei nur die Faktorladungen von Interesse. Das Resultat dieser Schätzung sind die Faktorladungen λ_{mj} , welche ein Maß dafür sind, wie stark die Antwort einer Person auf der Variablen x_{im} von dem jeweiligen Faktor beeinflusst wird.

Zur Durchführung einer Faktorenanalyse müssen bestimmte Voraussetzungen erfüllt sein. Nach Bühner (2006) müssen Ausreißerwerte und die Verteilung der Items, die Korrelationen zwischen den Items, die Stichprobengröße sowie die Itemzahl pro Faktor berücksichtigt werden.

Ausreißerwerte und die damit einhergehenden Abweichungen von der Normalverteilung können die Höhe der Korrelationen (und somit auch der Kovarianzen) zwischen Items beeinflussen. Dies kann die Ergebnisse einer Faktorenanalyse verzerren. Als Konsequenz muss vor der Durchführung einer Faktorenanalyse die Verteilung der Items geprüft werden.

Weiterhin müssen zwischen den Items ausreichend hohe Korrelationen bestehen, um überhaupt Faktoren extrahieren zu können. Um die Eignung von Daten zur Berechnung einer Faktorenanalyse zu bestimmen, gibt es mehrere Verfahren. Das am häufigsten benutzte ist das Kaiser-Meyer-Olkin-Maß (KMO-Maß), das den gemeinsamen Varianzanteil aller Variablen miteinander bestimmt und diesen mit den Summen des gemeinsamen Varianzanteils aller Variablen und des quadrierten Partialkorrelationskoeffizienten in Beziehung setzt (vgl. Bühner, 2006). Eine hohe Summe der Partialkorrelationskoeffizienten bedeutet, dass die Interkorrelationsmatrix wenig gemeinsame Varianz enthält, sodass der KMO-Koeffizient klein wird. Um die Eignung einer Interkorrelationsmatrix beurteilen zu können, gelten die in Tabelle 26 nach Kaiser und Rice (1974) zusammengefassten Anhaltspunkte, die in den nachfolgenden Untersuchungen genutzt werden.

Tabelle 27: Eignung der Daten zur Berechnung einer Faktorenanalyse mittels des KMO-Koeffizienten

KMO-Koeffizient	Eignung der Daten
> .90	sehr gut
.80 - .90	gut
.70 - .79	mittel
.60 - .69	mäßig
.50 - .59	schlecht
< .50	keine EFA berechenbar

Analog zum KMO-Maß, das die Eignung der gesamten Korrelationsmatrix zur Durchführung einer Faktorenanalyse angibt, lässt sich die Eignung eines Items zur Durchführung einer Faktorenanalyse bestimmen. Dieses Maß wird *Measure of Sampling Adequacy* (MSA) genannt und berechnet sich aus den Korrelationen und Partialkorrelationen zwischen einem Item und den anderen Items (Bühner, 2006). Das MSA-Maß gibt somit an, wie gut ein Item zu den übrigen Items passt, mit denen eine Faktorenanalyse durchgeführt werden soll. Die Interpretation des MSA richtet sich nach denselben Kriterien wie beim KMO-Maß.

Die Stichprobengröße und die Anzahl der Items beeinflussen ebenso die Ergebnisse einer Faktorenanalyse. In der Literatur zu findende Daumenregeln (siehe z. B. Brayant & Yarnold, 2001 für eine Übersicht) erwiesen sich als nicht hinreichend. MacCallum, Widaman, Zhang und Hong (1999) konnten in Monte-Carlo-Studien zeigen, dass nicht das Verhältnis von Items zu Probanden für die Ergebnisse entscheidend ist, sondern die Kommunalität (h^2) eines Items. Die Kommunalität gibt an, wie viel Varianz eines Items durch alle Faktoren erklärt wird. Nach MacCallum et al. (1999) ist eine Stichprobengröße von $n = 60$ ausreichend, sofern die Kommunalität der Items mindestens .60 beträgt. Ergebnisse von Mundform, Shaw und Ke (2005) bestätigen dies. Allerdings wurden die Ergebnisse von Mundform et al. (2005) nur an Maximum-Likelihood-Faktorenanalysen mit Varimax-Rotation (s. u.) gewonnen, sodass diese nicht auf andere Extraktions- und Rotationsverfahren verallgemeinert werden können. Zusammenfassend lässt sich der Zusammenhang von Stichprobengröße so darstellen, wie es in Tabelle 28 gezeigt wird (in Anlehnung an Bühner, 2006).

Tabelle 28: Beurteilung der Stichprobengröße für eine Faktorenanalyse

Stichprobengröße	Kommunalität	Beurteilung
$n < 60$	$h^2 < .60$	keine Faktorenanalyse durchführbar
$n = 60$	$h^2 > .60$	gerade ausreichend
$n = 100$	$h^2 > .50$	ausreichend
$n = 200$	$h^2 > .50$	fair
$n = 300$	$h^2 > .50$	gut
$n = 500$	$h^2 > .50$	sehr gut
$n = 1000$	$h^2 > .50$	exzellent

Neben der Inspektion der Itemverteilungen werden für die hier durchgeführten Faktorenanalysen in erster Linie das KMO-Maß sowie ggf. die Stichprobengröße im Verhältnis zu den Kommunalitäten zur Beurteilung der Eignung der Daten für die Durchführung einer Faktorenanalyse herangezogen. Die Betrachtung der Kommunalitäten wird nur dann durchgeführt, wenn eine geringe Stichprobengröße vorliegt.

Um die Faktorladungen schätzen zu können, stehen verschiedene Extraktionsverfahren zur Verfügung. Exemplarisch seien hier die Hauptkomponentenanalyse, die Hauptachsenanalyse sowie die Maximum-Likelihood-Faktorenanalyse genannt. Bei der Hauptkomponentenanalyse handelt es sich nicht um eine Faktorenanalyse im eigentlichen Sinn, da hier der additive Term mit der Residualvariablen nicht berücksichtigt wird. Hauptkomponentenanalysen stellen daher lediglich eine Zusammenfassung von Variablen dar, die in einem gewissen Maß korrelieren. Die eigentliche Faktorenanalyse will aber die Frage beantworten, welche latenten Variablen die manifesten Variablen kausal beeinflussen (Backhaus, Erichson, Blinke & Weiber, 2003).

Die Hauptachsenanalyse (PAF, *principal axis factor analysis*) sucht nach den latenten Faktoren, welche die Korrelationen zwischen den Items ursächlich erklären und bildet das *Common-factor*-Modell ab. Diese Art der Faktorenanalyse entspricht der Interpretation als reflexivem Messmodell. Das Besondere bei der PAF ist, dass bei der Analyse nur die Zusammenhänge mit anderen Items berücksichtigt werden, so dass niemals die gesamte Varianz reproduziert werden kann. Bei der Maximum-Likelihood-Faktorenanalyse (ML, *maximum likelihood factor analysis*) wird aus der Stichprobenkorrelationsmatrix zuerst die Populationskorrelationsmatrix geschätzt und anschließend werden aus dieser wie bei der PAF die Faktoren extrahiert. Das besondere an der ML-Faktorenanalyse ist, dass mittels eines χ^2 -Wertes ein Modelltest durchgeführt werden kann, der prüft, ob die extrahierte Faktorenstruktur den Daten der Population entspricht. Die Interpretation der Ergebnisse bei einer ML-Faktorenanalyse erweitert sich noch dahingehend, dass nicht nur die Frage beantwortet wird, welche latenten Variablen die Ausprägungen der manifesten Variablen beeinflussen, sondern ob diese Verursachung auch in der zugrunde liegenden Population gilt oder nicht.

Bei der Extraktion der Faktoren stellt sich außerdem die Frage, wie viele Faktoren extrahiert werden sollen. Die Zahl der zu extrahierenden Faktoren kann durch theoretische Vorüberlegungen oder anhand empirischer Kriterien entschieden werden. Die bekanntesten dieser Kriterien sind das Eigenwert-Kriterium, der Scree-Test nach Cattell (1966) sowie die Parallelanalyse nach Horn (1995). Bei dem Eigenwert-Kriterium werden so viele Faktoren extrahiert, wie Eigenwerte > 1 existieren. Der Eigenwert eines Faktors gibt an, wie viel Varianz durch einen Faktor ausgeklärt wird (Bortz, 1999). Ein Eigenwert > 1 bedeutet, dass ein Faktor mehr Varianz aufklärt als eine einzelne Variable. Das Kriterium besagt demnach, dass so viele Faktoren extrahiert werden, wie es Faktoren gibt, die mehr Varianz aufklären als ein Item.

Ein weiteres Kriterium zur Bestimmung der Anzahl der zu extrahierenden Faktoren ist der Scree-Test nach Cattell (1966). Beim Scree-Test werden zuerst die Eigenwerte in einem Diagramm dargestellt. Dabei werden die Nummer des Faktors auf der Abszisse und die dazugehörigen Eigenwerte auf der Ordinate abgetragen und danach die Punkte mit einer Geraden verbunden. Diese Darstellung wird Scree-Plot genannt. Die Stelle, an dem diese Linie einen Knick aufweist, bestimmt die Zahl der zu extrahierenden Faktoren. Die Zahl wird auf der Abszisse abgelesen. An diesem Punkt büßt die Methode ihre Objektivität ein. So weist Bühner (2004) darauf hin, dass in der einschlägigen Literatur sowohl Versionen des Scree-Tests zu finden sind, die nur die Zahl der Faktoren bis zum Knick als bedeutsam erachten, wohingegen nach anderen Angaben aber auch der Faktor nach dem Knick mitzählt. Auch kann es passieren, dass in einem Scree-Plot mehrere Knicke vorkommen. Somit hängt die Zahl der Faktoren von den Vorlieben des Auswerters ab.

Bei der Parallelanalyse werden Eigenwerte einer Faktoranalyse eines empirisch gewonnenen Datensatzes mit den Eigenwerten einer Faktorenanalyse eines Datensatzes mit Zufallsdaten verglichen (wobei diese Zufallsvariablen normal verteilt sind). Die Idee ist nun, dass bei der Faktorenanalyse des empirischen Datensatzes so viele Faktoren beibehalten werden, wie es Eigenwerte gibt, die größer als die Eigenwerte des Zufallsdatensatzes sind. Voraussetzung ist die Übereinstimmung der Anzahl der Variablen und der Fälle des empirischen Datensatzes und des Zufallsdatensatzes. Anders ausgedrückt heißt dies, dass die Faktoren, die man beibehalten will, wenigstens soviel Varianz aufklären sollen, wie sich bei der Analyse von Zufallsdaten ergibt. Allerdings gilt als übergeordnetes Kriterium bei der Bestimmung der Faktorzahl immer, dass die Faktorenlösung inhaltlich sinnvoll interpretierbar sein muss (Bühner, 2006).

Nach der Extraktion der Faktoren stellt sich das Problem, die gefundene Faktorlösung zu interpretieren. Da die Extraktion ausschließlich nach einem Algorithmus erfolgt, ist die Interpretierbarkeit der Lösung nicht immer gegeben. Um die Interpretierbarkeit zu erleichtern, sind Rotationsverfahren entwickelt worden, mit deren Hilfe die gefundenen Lösungen im Faktorraum so rotiert werden können, dass die Interpretation der Faktorladungsmatrix vereinfacht wird. Grundsätzlich unterscheidet man zwischen orthogonalen und obliquen Rotationsverfahren (Eckey, Kosfeld & Rengers, 2002). Bei der orthogonalen Rotation werden die Faktoren so rotiert, dass diese nicht korrelieren dürfen, d. h. geometrisch gesehen orthogonal aufeinander stehen. Die Rotation wird immer unter der Beachtung eines Rotationskriteriums durchgeführt. Bei einer orthogonalen Rotation ist das bekannteste Kriterium das Varimax-Kriterium (Eckey,

Kosfeld & Rengers, 2002). Bei diesem werden die Faktoren so rotiert, dass diese nacheinander möglichst viel Varianz aufklären, d. h. der erste Faktor erklärt nach der Rotation den größten Anteil der Varianz, der zweite Faktor den zweitgrößten Anteil usw. Grundsätzlich sind bei orthogonalen Rotationen die Faktorladungen identisch mit den Korrelationen der Items mit den Faktoren. Ziel der Varimax-Rotation ist die Einfachstruktur, d. h. die Items sollen nur hoch auf einem Faktor und niedrig auf allen anderen Faktoren laden (Bortz, 1999).

Neben der orthogonalen Rotation gibt es noch oblique Rotationsverfahren. Bei diesen dürfen die Faktoren korrelieren, d. h. geometrisch gesehen sind diese nicht mehr orthogonal. Der Vorteil von obliquen Rotationen gegenüber orthogonalen ist, dass der Zusammenhang von Faktoren bestimmt werden kann, ohne dass von vornherein die Unabhängigkeit der Faktoren festgelegt wird. Ist also nicht bekannt, ob die latenten Variablen zusammenhängen oder nicht, empfiehlt sich zuerst die Anwendung einer obliquen Rotationstechnik (Bühner, 2006). Sind die Korrelationen der Faktoren nur gering ausgeprägt, kann anschließend ein orthogonales Rotationsverfahren angewendet werden. Ähnlich wie bei der orthogonalen Rotation gibt es auch für die oblique Rotation bestimmte Rotationsverfahren, die jeweils nach einem anderen Kriterium vorgehen. Das bekannteste der obliquen Rotationsverfahren ist das Promax-Verfahren. Bei diesem werden die ursprünglichen orthogonalen Ladungen mit den Exponenten 2, 4 oder 6 potenziert und anschließend oblique rotiert (Eckey, Kosfeld & Rengers, 2002). Dadurch werden extrem hohe und extrem niedrige Faktorladungen vermieden und die Interpretation der Ergebnisse wird erleichtert. Die Exponenten müssen vor der Durchführung der Rotation festgelegt werden. Das Promax-Verfahren mit dem Exponenten 4 führt in der Regel zu guten Lösungen (Bühner, 2006, Eckey, Kosfeld & Rengers, 2002).

Anders als bei der orthogonalen Rotation kann nach einer obliquen nicht mehr die Kommunalität (die Summe der quadrierten Faktorladungen) interpretiert werden, da die Faktoren nicht mehr unabhängig voneinander sind und somit einen gemeinsamen Varianzanteil aufweisen (Eckey, Kosfeld & Rengers, 2002). Ein weiterer Unterschied ist, dass bei obliquen Rotationsverfahren die ursprüngliche Faktorladungsmatrix in eine Mustermatrix und eine Strukturmatrix aufgeteilt wird. Die Mustermatrix gibt die standardisierten partiellen Regressionsgewichte der Items auf die Faktoren an, während die Strukturmatrix die Korrelationen der Items und der Faktoren angibt. Zur Interpretation sollte die Mustermatrix betrachtet werden (Bühner, 2006). Weiterhin wird bei obliquen Rotationsverfahren die Matrix der Faktorkorrelation ausgegeben.

An dieser Stelle muss auf ein grundlegendes, aber selten diskutiertes Merkmal der Faktorenanalyse eingegangen werden. Begonnen bei der Entscheidung, welche Variablen in die Faktorenanalyse aufgenommen werden, über die Entscheidung der Anzahl der zu extrahierenden Faktoren bis zur Auswahl der Rotationsverfahren sind Ergebnisse einer Faktorenanalyse wesentlich durch die zuvor genannten (und teilweise arbiträren) Schritte bestimmt. Bei einer Faktorenanalyse lässt sich auch kein Faktor finden, für den im Vorhinein keine Indikatoren aufgenommen wurden. Wird als Rotationsmethode eine orthogonale Rotationstechnik gewählt, so sind die resultierenden Faktoren immer unabhängig voneinander, egal ob dies der Realität entspricht oder nicht. Auch bei einer obliquen Rotationstechnik spiegeln die ermittelten Faktorkorrelationen nicht die Realität wider, weil wie z. B. bei der Promax-Rotation der Exponent festgelegt werden muss, der maßgeblich die Ergebnisse der Rotation bestimmt. Eine erste Folge ist, dass jeder Schritt bei der Durchführung einer Faktorenanalyse demnach theoretisch oder empirisch begründbar sein muss und dokumentiert werden sollte. Eine zweite Folgerung ist die Kontextabhängigkeit faktorenanalytischer Ergebnisse. So hat Beauducel (2001) mithilfe einer Simulationsstudie gezeigt, dass die Ergebnisse einer Hauptachsenanalyse von der Ausprägtheit einer Einfachstruktur in einem künstlich erzeugten Datensatz abhängen. Bei der Hinzunahme weiterer Variablen, die nur wenig mit den eigentlichen Indikatorvariablen eines Faktors korrelierten, zeigte sich, dass die ursprünglichen Indikatorvariablen nicht mehr eindeutig einem Faktor zuzuordnen waren und auch die Anzahl der zu extrahierenden Faktoren als Ergebnis einer Parallelanalyse von der wirklich zugrunde liegenden Faktorenanzahl abwich. Entscheidend für die Ergebnisse ist also immer auch, welche Items aufgenommen werden. Auch hier sollte die Entscheidung über die Aufnahme bestimmter Items aufgrund theoretischer Überlegungen erfolgen.

Bei der Anwendung der explorativen Faktorenanalyse bei den hier durchgeführten Studien wird das folgende Vorgehen angewandt: Zuerst wird mithilfe des KMO-Maßes und – wenn erforderlich – mithilfe der Richtwerte aus Tabelle 28 für die Beurteilung der Kommunalitäten überprüft, ob die Daten für die Durchführung einer Faktorenanalyse geeignet sind. Als Extraktionsalgorithmus wird die Hauptachsenanalyse bzw. die Maximum-Likelihood-Faktorenanalyse angewandt, da diese dem Modell gemeinsamer Faktoren und damit der gängigen Vorstellung latenter Variablen entsprechen. Die Anzahl der zu extrahierenden Faktoren wird nach theoretischen Gesichtspunkten bzw. empirischen Befunden aus der gängigen Literatur festgelegt. Die explorative Faktorenanalyse wird im Rahmen der durchgeführten Studien dann eingesetzt, wenn nach faktoriellen Strukturen gesucht werden soll. Dies spiegelt den

Charakter der EFA als sturkturentdeckendes Verfahren wider. Als Rotationsverfahren wird bei explorativen Faktorenanalysen die Promax-Rotation verwendet. Neben der Möglichkeit auf explorativem Weg nach Strukturen zu suchen, gibt es zudem die Möglichkeit, vorhandene oder postulierte faktorielle Strukturen auf ihre Passung empirischen Daten hin zu prüfen. Hierzu wird die konfirmatorische Faktorenanalyse angewandt, die im folgenden Kapitel beschrieben wird.

3.2 Konfirmatorische Faktorenanalyse

Das Verfahren der konfirmatorischen Faktorenanalyse (CFA, *confirmatory factor analysis*) stammt, im Gegensatz zur explorativen Faktorenanalyse, aus der Gruppe der hypothesenprüfenden Verfahren. Bei der CFA wird nicht nach einer unbekanntem Struktur in den Daten gesucht, sondern es wird überprüft, ob ein theoretisches Modell die Daten angemessen repräsentiert. Das heißt, die Zuordnung von latenten Variablen und manifesten Variablen erfolgt auf der Grundlage einer Theorie, und die CFA prüft, ob zwischen diesem Modell und den vorliegenden Daten eine Passung – Modellfit oder auch Modellpassung genannt – besteht (Schermelleh-Engel & Moosbrugger, 2008). Ist der Modellfit, der anhand bestimmter Parameter überprüft wird, hinreichend hoch, kann davon ausgegangen werden, dass das theoretische Modell die Daten angemessen repräsentiert.

3.2.1 Grundlegende CFA-Modelle

Die Grundannahme der CFA ist exakt wie bei der EFA das Modell gemeinsamer Faktoren nach Thurstone (1947). Grafisch lässt sich ein Modell mit einer latenten und drei manifesten Variablen wie in Abbildung 12 veranschaulichen, wobei eine spezielle Symbolik genutzt wird. Die manifesten Variablen werden grafisch immer durch ein Quadrat (\square) und die latenten Variablen durch einen Kreis (\circ) dargestellt. Korrelationen zwischen Variablen werden durch einen Doppelpfeil (\leftrightarrow) symbolisiert. Pfeile (\rightarrow) stellen allgemein ein Regressionsgewicht dar, wobei im Kontext von CFA-Modellen das Regressionsgewicht auch als Faktorladung bezeichnet wird. Werden diese Parameter in einem Modell nicht frei geschätzt (sogenannte fixierte oder restringierte Parameter, vgl. Fußnote 33), werden die festgesetzten Parameterwerte neben das jeweilige Symbol geschrieben. So bedeutet z. B. der Wert 1 über dem Regressionsgewicht der Fehlervariablen, dass dieses Regressionsgewicht auf den Wert 1 festgelegt wurde. Das in Abbildung 12 dargestellte Modell entspricht einem direkt-reflexiven Messmodell (Edwards & Bagozzi, 2000).

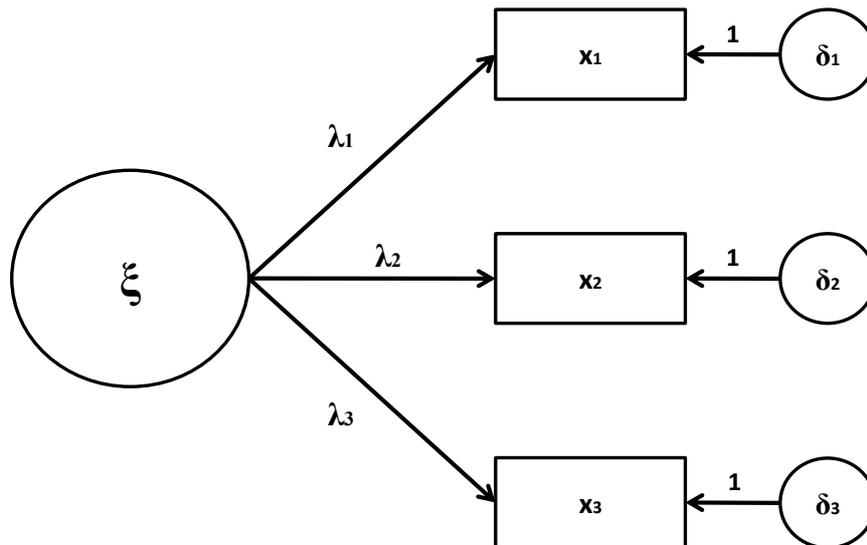


Abbildung 12: Darstellung eines CFA-Modells mit einer latenten Variablen

Für den Fall mit zwei latenten Variablen lässt sich das CFA-Modell grafisch wie in Abbildung 13 darstellen.

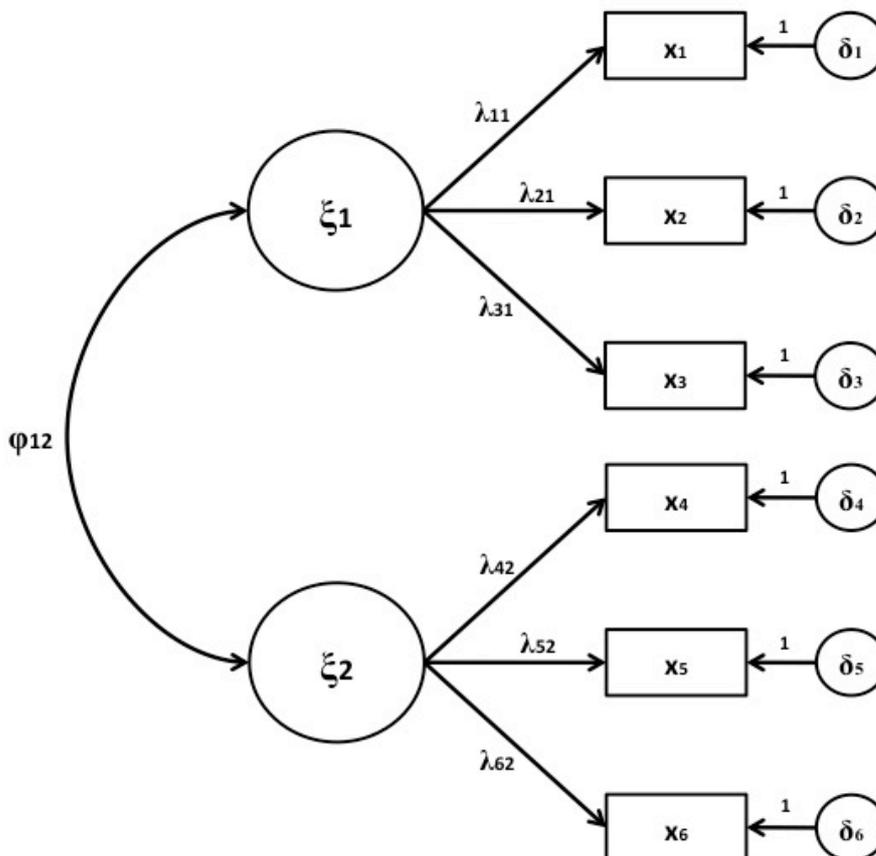


Abbildung 13: Darstellung eines CFA-Modells mit zwei latenten Variablen

Bei einer Korrelation der latenten Variablen werden alle manifesten Variablen aufgrund der Korrelation indirekt durch alle latenten Variablen beeinflusst. Dabei wird eine latente Variable ξ_1 direkt durch die ihr zugeordneten manifesten Variablen x_{11} , x_{12} , x_{13} beeinflusst. Die la-

tente Variable ξ_2 beeinflusst die manifesten Variablen x_{11} , x_{12} , x_{13} indirekt über die Korrelation ϕ_{12} mit der latenten Variablen ξ_1 . Die Stärke der indirekten Wirkung ist somit auch abhängig von der Größe der Korrelation ϕ_{12} . Dies gilt natürlich auch entsprechend für eine indirekte Wirkung der latenten Variablen ξ_1 auf die manifesten Variablen x_{21} , x_{22} , x_{23} .

Diese beiden Modelle stellen die grundlegenden Modelle für eine CFA dar. Hierzu können nun noch mehr Items bzw. mehr latente Variablen in das Modell aufgenommen werden. Werden latente Variablen in das Modell aufgenommen, erhöht sich auch die Anzahl der Korrelationen zwischen diesen Variablen, die modelliert werden können. In einer CFA kann vorgegeben werden, ob die Korrelation zwischen latenten Variablen in das Modell aufgenommen wird oder nicht. Wird keine Korrelation angenommen, sind die latenten Variablen unabhängig voneinander, in der Terminologie der EFA kann dann von orthogonalen Faktoren gesprochen werden.

Bislang wurden nur CFA-Modelle vorgestellt, bei denen die manifesten Variablen durch jeweils eine latente Variable direkt bzw. indirekt über die Korrelation ϕ_{12} beeinflusst wurden. Alternativ können aber auch Kreuzladungen modelliert werden. Eine Kreuzladung bedeutet, dass eine manifeste Variable nicht nur wie zuvor beschrieben direkt durch eine latente Variable beeinflusst wird, sondern durch zwei oder mehr. Dies wird in Abbildung 14 illustriert, in der die manifesten Variablen x_1 bis x_6 durch die beiden latenten Variablen ξ_1 und ξ_2 beeinflusst werden.

Die Beachtung aller möglichen Kreuzladungen stellt eine Möglichkeit dar, die CFA mit der EFA zu verknüpfen. Wie oben dargestellt, wird bei der EFA davon ausgegangen, dass eine manifeste Variable durch alle latenten Variablen beeinflusst wird. Durch die Modellierung aller möglichen Kreuzladungen wird also das EFA-Modell abgebildet. Dies bietet die Möglichkeit, ein mittels EFA ermitteltes Faktormodell mit den Möglichkeiten der CFA empirisch zu prüfen. Dieses Vorgehen wird E/CFA genannt (Brown, 2006) und stellt eine Modellierung einer EFA mithilfe einer CFA dar. Insbesondere erlaubt es die genauere Untersuchung einer auf explorativem Weg ermittelten Faktorenstruktur. So lassen sich bei der E/CFA z. B. die Messfehler untersuchen, die bei einer EFA nicht berechnet werden oder es lassen sich die Korrelationen zwischen den Faktoren (bzw. latenten Variablen) ermitteln, ohne die Ein-

schränkungen der orthogonalen oder obliquen Rotationsverfahren hinzunehmen, und überdies lässt sich für ein E/CFA-Modell die Passung testen (vgl. Brown, 2006).²⁹

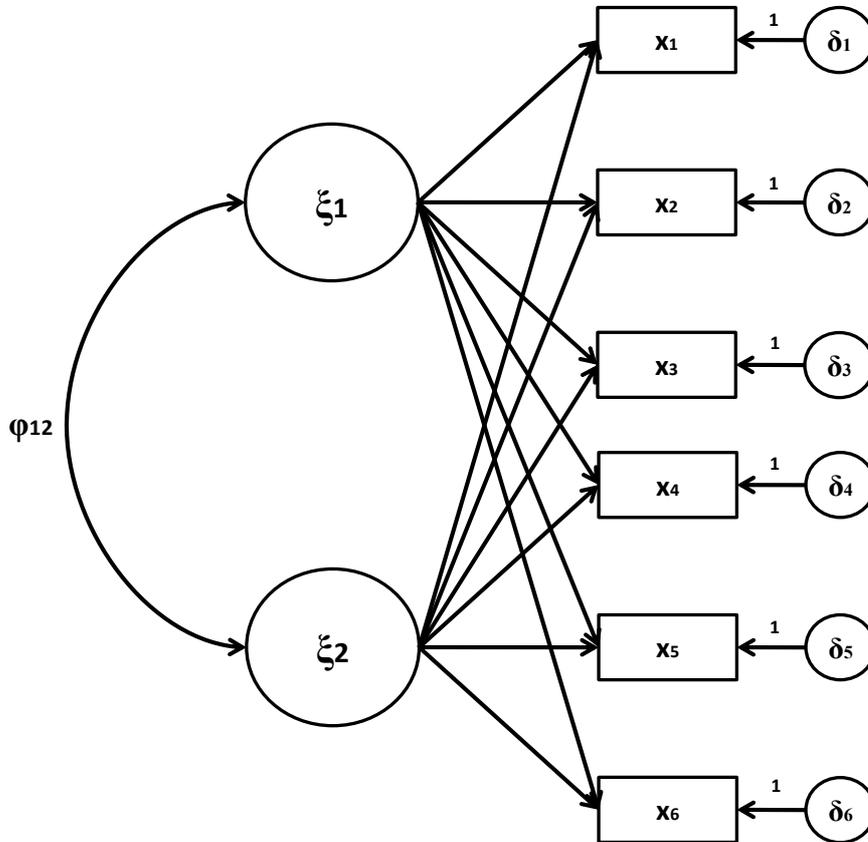


Abbildung 14: CFA-Modell mit zwei latenten Variablen und Kreuzladungen zu allen manifesten Variablen

Die vorgestellten CFA-Modelle gehen von unkorrelierten Fehlervarianzen aus. Es ist aber auch möglich, die Korrelation von Messfehlern in einem CFA-Modell zu modellieren, Abbildung 15 stellt dies grafisch durch die angenommene Korrelation der Messfehler der manifesten Variablen x_2 und x_3 dar. Die Annahme korrelierter Messfehler wird z. B. dann benötigt, wenn es bei der Messung der manifesten Variablen zu Methodeneffekten kommt (vgl. Schermelleh-Engel & Schweizer, 2008). Im Rahmen der vorliegenden Arbeit werden korrelierende Messfehler zur Beschreibung von Effekten der Itemformulierung gebraucht (Schriesheim, Eisenbach & Hill, 1991, DiStefano & Motl, 2009). Korrelierte Messfehler treten häufig dann auf, wenn z. B. Items gleiche Wörter enthalten und ähnlich formuliert sind bzw. wenn in einer Skala Items vorhanden sind, die positiv und negativ formuliert sind. Korrelierende Messfehler entsprechen Methodenfaktoren, die systematische Varianzquellen repräsentieren, die neben

²⁹ Eine Ausnahme hier ist die Maximum-Likelihood-Faktorenanalyse, bei der ein χ^2 -Wert zur Bestimmung der Modellpassung berechnet wird. Darüber hinaus werden aber keine weiteren Fitindizes berechnet und die oben beschriebenen Nachteile einer EFA finden sich auch bei der Maximum-Likelihood-Faktorenanalyse.

der eigentlichen latenten Variablen die Itembeantwortung beeinflussen (vgl. Schermelleh-Engel & Schweizer, 2008).

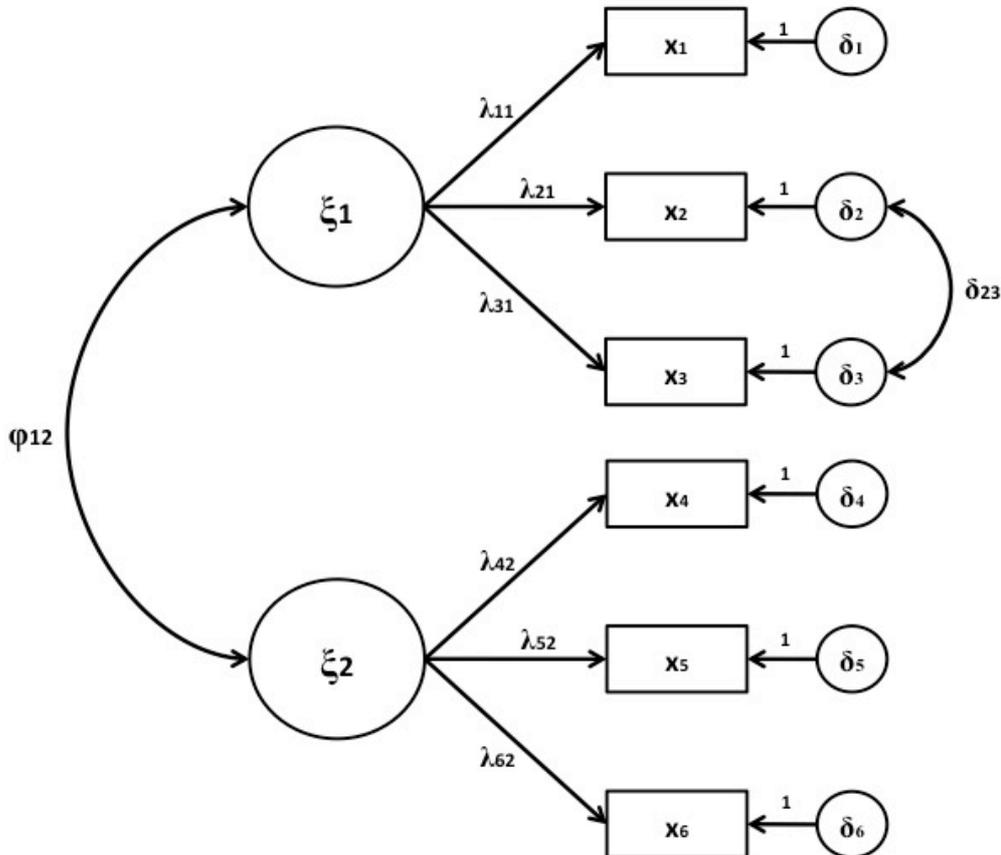


Abbildung 15: CFA-Modell mit korrelierten Messfehlern

Nachdem ein CFA-Modell aufgestellt wurde, muss daran anschließend getestet werden, ob das Modell die empirischen Daten richtig beschreibt. Hierbei wird die Hypothese geprüft, ob das Modell zu den empirischen Daten passt oder nicht. Die zu prüfenden Hypothesen lauten (Bühner, 2006):

H_0 : Das Modell passt zur Datenstruktur.

H_1 : Das Modell weicht von der Datenstruktur ab.

Zur Testung des Modells kann ein χ^2 -Wert herangezogen werden. Der χ^2 -Wert errechnet sich aus der Maximum-Likelihood-Schätzung einer Fitfunktion, die sich aus der Differenz zwischen den durch das Modell implizierten Zusammenhängen der manifesten Variablen und deren empirisch gefundenen Zusammenhängen ableiten lässt (vgl. Brown, 2006, für eine ausführliche Beschreibung). Der aus der Fitfunktion abgeleitete χ^2 -Wert besitzt $[p(p+1)/2]-q$ Freiheitsgrade, wobei p die Anzahl der manifesten Variablen im Modell angibt und q die Anzahl der zu schätzen Parameter darstellt.

Bei der Modellprüfung bedeutet ein nicht-signifikanter χ^2 -Wert, dass das theoretisch angenommene Modell zutrifft. Dies ist der üblichen Richtung bei der Testung von Hypothesen entgegengesetzt. In diesem Fall ist nicht der α -Fehler möglichst klein zu halten, sondern der β -Fehler (Bühner, 2006). Die Problematik hierbei ist jedoch, dass der β -Fehler nur dann geschätzt werden kann, wenn die Abweichung der Verteilungen unter der Annahme der Null- bzw. der Alternativhypothese bekannt ist. In der Regel ist aber nur die Verteilung unter Annahme der Nullhypothese bekannt. Der β -Fehler muss somit durch die Erhöhung des gewählten Signifikanzniveaus „künstlich verkleinert“ werden.³⁰ Allerdings wird dadurch der exakte Wert immer noch nicht bekannt. In der Regel setzt man für die Prüfung solcher Hypothese ein $\alpha \geq .20$ fest.

Die Maximum-Likelihood-Schätzung der Fitfunktion stellt Anforderungen an die Verteilung der Daten. So werden die Normalverteilung der Daten sowie eine ausreichend große Stichprobengröße gefordert. Die Maximum-Likelihood-Methode liefert aber auch bei moderaten Verletzungen der Normalverteilung und kleineren Stichprobengrößen ab $n \geq 100$ sehr gute Ergebnisse (Schermelleh-Engel, Moosbrugger & Müller, 2003). Allerdings können bei kleinen Stichproben Schätzprobleme auftauchen (Marsh, Hau, Balla & Grayson, 1998), sodass als Mindestgröße für eine CFA eine Stichprobengröße von 200 zu empfehlen ist (Bühner, 2006). Zur Überprüfung der Normalverteilungsannahmen können die einzelnen Verteilungen der Variablen direkt untersucht werden oder es kann der Mardia-Test auf multivariate Normalverteilung durchgeführt werden (Bühner, 2006). Bezüglich der geforderten Normalverteilung sind Abweichungen bis zum Absolutbetrag der Schiefe ≤ 2 und dem Absolutbetrag der Wölbung ≤ 7 tolerierbar, solange die Stichprobengröße $n \geq 100$ ist (West, Finch & Curran, 1995). Bei den Angaben von West et al. (1995) ist allerdings zu berücksichtigen, dass der Begriff der Wölbung nicht das vierte Moment der Verteilung (vgl. Bortz, 1999) bezeichnet, sondern die Differenz des vierten Moments minus 3 gemeint ist. Diese Differenz wird auch als Exzess einer Verteilung bezeichnet (Bortz, 1999, West et al., 1995). Ihre Verwendung rührt daher, dass eine Normalverteilung eine Wölbung von 3 hat und der Exzess somit die Abweichung der Wölbung einer gegebenen Verteilung von einer Normalverteilung wiedergibt.³¹

³⁰ Dies resultiert aus dem allgemeinen Zusammenhang, wonach der β -Fehler kleiner wird, wenn das gewählte Signifikanzniveau erhöht wird (vgl. Bortz, 1999).

³¹ Der Exzess wird auch in der in hier eingesetzten Software AMOS automatisch ausgegeben, d. h. die später berichteten Wölbungen entsprechen dem Exzess der Verteilung.

Dennoch kann es auch innerhalb der genannten Grenzen zu Verzerrungen kommen. So wird sehr leicht der χ^2 -Wert überschätzt, was zu einer vorschnellen Ablehnung des Modells führen kann. Zu diesem Zweck wird der Bollen-Stine-Bootstrap eingesetzt, bei dem die Wahrscheinlichkeit des χ^2 -Wertes korrigiert wird (Bollen & Stine, 1991). Auch wird bei Abweichungen von der Normalverteilung die Schätzung der Standardfehler der Modellparameter fälschlicherweise kleiner, sodass bei der Signifikanztestung diese eher signifikant werden. Für die Auswertung wird daher immer die Verteilungsform der Daten geprüft und bei auffälligen Abweichungen von der Normalverteilung bzw. einem signifikanten Mardia-Test der Bollen-Stine-Bootstrap angewandt. Bei der Interpretation der Parameter muss dann die Möglichkeit berücksichtigt werden, dass deren Signifikanz zu hoch ausfällt.

Ein Modell kann aber nur auf Signifikanz geprüft werden, wenn die Anzahl der Freiheitsgrade der χ^2 -Verteilung größer null ist. Hier stellt sich die Frage der Modellidentifikation. Wie aus dem Freiheitsgrad des χ^2 -Werts deutlich wird, hängt die Modellidentifikation von der Anzahl der nicht-redundanten Elemente in der empirischen Korrelationsmatrix sowie von der Anzahl der zu schätzenden Parameter ab. Die nicht redundanten Elemente der empirischen Kovarianzmatrix sind die Elemente, die in der unteren bzw. oberen Dreiecksmatrix zu finden sind. Dort stehen die Varianzen und Kovarianzen der manifesten Variablen, welche die zur Berechnung des entstehenden Gleichungssystems notwendige Information liefern. Die Anzahl der nicht redundanten Elemente der empirischen Kovarianzmatrix lässt sich nach Brown (2006) mithilfe des Ausdrucks $b = p(p+1)/2$ berechnen.

Für den Fall von drei manifesten Variablen liegen also sechs nicht redundante Elemente vor. Nun können sich drei Fälle ergeben (q bezeichnet die Anzahl der zu schätzenden Parameter):

- $b < q$: In diesem Fall liegen nicht genügend Informationen vor, um die Modellparameter zu schätzen, man sagt, das Modell ist nicht identifiziert.
- $b = q$: Es liegen genauso viele Informationen vor, wie es Parameter zum Schätzen gibt. Man sagt, das Modell ist genau identifiziert. Bei genau identifizierten Modellen ist aber keine Signifikanzprüfung möglich, da die Freiheitsgrade der χ^2 -Verteilung gleich null sind.
- $b > q$: In diesem Fall ist genügend Information vorhanden, um alle Modellparameter zu schätzen, und die Freiheitsgrade der χ^2 -Verteilung sind größer null. Man sagt, das Modell ist überidentifiziert und es kann die Signifikanz des χ^2 -Wertes ermittelt werden.

Von Interesse ist folglich nur ein überidentifiziertes Modell, da nur bei diesem eine Signifikanzprüfung vorgenommen werden kann. Wie ein Beispiel eines Modells mit drei manifesten Variablen (siehe Abbildung 12) zeigt, sind hier insgesamt sieben Parameter zu berechnen, wohingegen nur sechs Informationseinheiten vorliegen. Daraus folgt, dass dieses Modell nicht ohne weitere Einschränkungen zu berechnen ist.

Hier kommt der Sachverhalt zum Tragen, dass die Metrik der latenten Variablen bestimmt werden muss (Kline, 1998). Dies wird dadurch erreicht, dass man entweder den Parameter λ einer manifesten Variablen oder die Varianz der latenten Variablen auf den Wert 1 fixiert. Welcher Wert fixiert wird, ändert nichts an den Schätzungen der übrigen Parameter oder an dem Fit des Gesamtmodells (Bühner, 2006). In der Regel wird immer das erste Item einer Skala (in der Reihenfolge, wie diese von den Autoren vorgegeben wurde) als Skaliervariable herangezogen. Wenn dieses Item aus irgendwelchen Gründen aus der Skala genommen wird, folgt das zweite Item. Sind in dem theoretischen Modell Restriktionen vorgegeben, welche z. B. eine Gleichsetzung aller Parameter λ_i voraussetzen, wird die Varianz der latenten Variablen fixiert.

Wegen der Notwendigkeit, die Metrik der latenten Variablen festzulegen, fällt ein zu schätzender Parameter weg. Dadurch wird das zuvor genannte Beispiel mit drei manifesten Variablen zu einem gerade identifizierten Modell. Nach Kline (1998) sind CFA-Modelle immer genau dann identifiziert, wenn drei manifeste Variablen vorliegen. Allerdings ist dann noch keine Signifikanzprüfung möglich. Bei den verwendeten CFA-Modellen wird immer versucht, mindestens vier Items in einer Skala zu haben, um die Modellgüte angemessen prüfen zu können.

An dieser Stelle muss spezifiziert werden, welche Parameter bei der Schätzung eines CFA-Modells berechnet werden. Zum einen werden die im folgenden Kapitel beschriebenen Indizes zur Beurteilung der Modellpassung berechnet. Zum anderen werden, ähnlich wie bei einer EFA, die Faktorladungen λ_i des CFA-Modells berechnet. Weiterhin werden die Varianzen der latenten Variablen ξ und die Fehlervarianzen δ_i geschätzt. Sofern ein CFA-Modell mit mehr als zwei latenten Variablen vorliegt, wird ggf. deren Kovariation ϕ geschätzt. Die Parameter werden sowohl unstandardisiert als auch standardisiert geschätzt. Mittels der standardisierten Parameter erfolgt auch die Signifikanzprüfung der Parameter. Hierzu werden die *critical ratios* (CR) bestimmt, die das Verhältnis der Schätzwerte-Parameter zu der Standardabweichung

chung der Schätzwerte angeben. Das *critical ratio* ist damit analog einem z-Wert zu interpretieren, d. h. CR-Werte, deren Absolutbeträge über 1.96 liegen, weisen auf einen auf dem 5%-Niveau signifikanten Parameter hin, wobei eine zweiseitige Testung zugrunde gelegt wird. Entsprechend weisen CR-Werte, deren Absolutbeträge über dem Wert von 2.57 liegen, auf einen auf dem 1%-Niveau signifikanten Parameter hin.

Von besonderem Interesse ist aber, ob das theoretisch angenommene Modell zu den empirischen Daten passt. Dies lässt sich mit dem zuvor genannten χ^2 -Test überprüfen der aber nicht sehr zuverlässig ist, sodass noch weitere Indizes angeführt werden sollen, welche die Modellpassung beschreiben

3.2.2 Bestimmung der Modellpassung

Die Beurteilung der Modellanpassung stellt kein leichtes Unterfangen dar, weil keine durchgängig zu interpretierenden statistischen Tests vorhanden sind und zudem eine Vielzahl an Fitindizes existiert, von denen jeder einen besonderen Aspekt der Modellpassung beschreibt (Schermelleh-Engel et al., 2003). In diesem Kapitel wird zuerst der übliche χ^2 -Test mit seinen Mängeln beschrieben und die nachfolgend benutzen Fitindizes werden näher erläutert.

Der zuvor beschriebene χ^2 -Test ist der Kennwert, welcher sich direkt aus der Fitfunktion eines Modells errechnen lässt. Mithilfe der angegebenen Freiheitsgrade lässt sich die Wahrscheinlichkeit für den errechneten χ^2 -Wert bestimmen und somit eine Signifikanzprüfung des theoretischen Modells durchführen. Allerdings ist der χ^2 -Test mit Schwächen behaftet (vgl. zusammenfassend Schermelleh-Engel et al., 2003). Neben den schon dargestellten Schwierigkeiten beim Bestimmen des α -Niveaus kommt hinzu, dass der χ^2 -Wert direkt von der Stichprobengröße n abhängig ist. Dies führt dazu, dass aus einem steigenden Stichprobenumfang ein hoher χ^2 -Wert resultiert. Die Folge ist, dass große Stichproben leichter signifikante Ergebnisse hervorbringen und damit einhergehend auch nur minimale Abweichungen der empirischen Daten vom theoretischen Modell zu dessen Ablehnung führen. Ein weiterer Schwachpunkt ist die Normalverteilung der Daten, welche der χ^2 -Test an die Daten stellt. Abweichungen von der Normalverteilung führen zu verzerrten χ^2 -Werten, die aber mittels des zuvor genannten Bollen-Stine-Bootstraps verbessert werden können. Ein letzter Nachteil des χ^2 -Tests ist seine Abhängigkeit von der Modellkomplexität, die sich in den Freiheitsgraden widerspiegelt. In dem Kontext von CFA-Modellen (oder allgemeiner der Strukturgleichungsmodelle) wird die Modellkomplexität durch die Anzahl der manifesten Variablen im Modell angege-

ben. Dies kann zu einer hohen Anzahl an Freiheitsgraden führen, woraus leicht ein nicht signifikanter χ^2 -Wert resultiert (vgl. Bortz, 1999 für den Zusammenhang des χ^2 -Wertes und den Freiheitsgraden). Der herkömmliche χ^2 -Test ist also kein optimaler Weg, um die Passung eines Modells zu beurteilen.

Ein möglicher Ausweg ist der Vergleich des empirischen χ^2 -Wertes mit dem Erwartungswert der χ^2 -Verteilung. Der Erwartungswert des χ^2 -Wertes ist gleich dessen Freiheitsgrad (vgl. Schermelleh-Engel et al., 2003). Jöreskog und Sörbom (1993) haben dazu angeregt, den χ^2 -Wert nicht im Sinne eines Signifikanztests zu nutzen, sondern im Sinne eines deskriptiven Index. Dazu wird das Verhältnis von empirischem χ^2 -Wert und dessen Erwartungswert gebildet, wobei der Erwartungswert eines χ^2 -Wertes der Anzahl seiner Freiheitsgrade entspricht. Für das Verhältnis ergibt sich somit χ^2/df . Liegt ein guter Modellfit vor, ist die Abweichung des χ^2 -Wertes vom Erwartungswert gering, das Verhältnis sollte im Intervall $2 \leq \chi^2 \leq 3$ liegen. Schermelleh-Engel et al. (2003) empfehlen $0 \leq \chi^2 \leq 2$ für eine gute Modellpassung und $2 \leq \chi^2 \leq 3$ für eine akzeptable Modellpassung. Allerdings ist dieses Verhältnis immer noch abhängig von der Stichprobengröße (Bollen, 1989).

Die grundlegende Abhängigkeit des χ^2 -Wertes von der Stichprobengröße und der Modellkomplexität haben dazu geführt, dass eine Reihe deskriptiver Indizes entwickelt wurde, um die Modellpassung zu beurteilen. Diese Indizes heißen deskriptiv, weil die Verteilungen ihrer Werte unbekannt sind und somit keine statistische Absicherung der Werte möglich ist (Hu & Bentler, 1995). Grundsätzlich gibt es zwei Arten dieser Indizes, und zwar Goodness-of-fit-Indizes und Badness-of-fit-Indizes (Backhaus et al., 2003). Bei Goodness-of-fit-Indizes bedeutet ein hoher numerischer Wert, dass das theoretische Modell gut zu der Struktur der Daten passt. Bei den Badness-of-fit-Indizes bedeutet ein höherer numerischer Wert, dass das theoretische Modell sehr von der Struktur der Daten abweicht. Diese Art der Klassifikation ist aber nicht sehr verbreitet. Üblicherweise teilt man die Fitindizes in die folgenden Kategorien ein (Schermelleh-Engel et al., 2003):

- Overall-Model-Fit-Indizes geben an, inwiefern ein theoretisches Modell den empirischen Daten entspricht. Die Grundlage ihrer Berechnung ist die Differenz der modellimplizierten Kovarianzmatrix und der empirischen Kovarianzmatrix.
- Auf Modellvergleiche basierende Indizes beruhen auf dem Vergleich des Fits eines theoretischen Modells mit dem Fit eines Basismodells zu vergleichen.

Overall-Model-Fit-Indizes und auf Modellvergleiche basierende Indizes finden regelmäßig in der Beurteilung eines Modells Anwendung. Bei der Auswertung eines CFA-Modells sollten immer alle Arten von Indizes sowie der χ^2 -Wert berücksichtigt werden (McDonald & Ho, 2002).

Der erste zu beschreibende Index aus der Kategorie der Overall-Model-Fit-Indizes ist der RMSEA, was *Root Mean Square Error of Approximation* bedeutet und auf der Idee basiert, dass die oben aufgeführte Nullhypothese, das theoretische Modell passe exakt zur Datenstruktur, in praktischen Anwendungen fast immer falsch ist. Browne und Cudeck (1993) haben daher vorgeschlagen, die Hypothese der exakten Modellpassung mit der Hypothese zu ersetzen, dass das theoretische Modell approximativ zu der empirischen Datenstruktur passt. Der RMSEA ist nach unten beschränkt. Höhere Werte deuten auf eine schlechtere Modellpassung hin, der RMSEA gehört somit zu der Gruppe der Badness-of-fit-Indizes. Üblicherweise wird bei der Angabe des RMSEA auch das 90%-Konfidenzintervall angegeben. Das Konfidenzintervall sollte bei einem guten Fit den Wert 0 enthalten und möglichst eng um den RMSEA-Wert liegen. Das letzte Kriterium ist sehr unscharf definiert und wird daher nur beiläufig bei der Auswertung berücksichtigt. Es existiert für den RMSEA der Test auf *close fit*, der die Nullhypothese $RMSEA < .05$ gegen die Alternativhypothese $RMSEA \geq .05$ prüft. Die Wahrscheinlichkeit dieses Tests sollte die Signifikanzgrenze nicht überschreiten (wobei sich hier wieder die Problematik der Wahl der Signifikanzniveaus stellt, da hier die eigentlich zu testende Hypothese die Nullhypothese und nicht wie üblich die Alternativhypothese ist).

Der nächste verwendete Fitindex auf der Kategorie der Overall-Fit-Indizes ist der SRMR, was eine Abkürzung für *Standardized Root Mean Square Residual* ist. Dieser Index beruht auf der Idee, die Passung eines Modells mithilfe der standardisierten Residuen zu beurteilen. Diese geben die standardisierten Abweichungen der empirischen Kovarianzmatrix von der modellimplizierten Kovarianzmatrix an. Ein Wert des SRMR von 0 gibt an, dass die empirische und die modellimplizierte Kovarianzmatrix übereinstimmen und das Modell sehr gut zu den Daten passt. Je größer der SRMR wird, desto größer ist die Abweichung der Matrizen und desto schlechter passt das theoretische Modell zu den Daten. Höhere Werte deuten auf eine schlechtere Modellpassung hin, der SRMR gehört somit zu der Gruppe der Badness-of-fit-Indizes.

Damit ist die Beschreibung der Overall-Fit-Indizes abgeschlossen und es folgt die Beschreibung der Indizes, die auf einem Modellvergleich basieren. Dazu ist es wichtig, die Modelle zu

kennen, mit denen ein Vergleich stattfindet. Jöreskog und Sörbom (1993) unterscheiden dazu folgende Modelle: Das *default model* gibt das jeweils spezifizierte Modell an. Das *independence model*, das zugleich als Basis des Vergleichs dient, ist ein sehr restriktives Modell. Beim *independence model* werden die Faktorladungen auf 1 und die Fehlervariablen auf 0 fixiert, zudem wird angenommen, dass alle Variablen unkorreliert sind. Das *independence model* hat somit einen schlechteren Fit und dient als Grundlage des Vergleichs. Das *saturated model* hingegen reproduziert exakt die beobachtete Kovarianzmatrix der Stichprobe und hat somit einen exakten Fit (Bühner, 2006).

Der erste zu beschreibende Index ist der *Tucker-Lewis-Index* (TLI; vgl. Bentler & Bonnett, 1980). Beim TLI wird die Abweichung des *default model* vom *independence model* betrachtet. Der TLI kann Werte zwischen 0 und 1 annehmen, ist aber nicht normiert, sodass auch Werte über 1 möglich sind. Höhere Werte deuten auf eine bessere Modellpassung hin, der TLI gehört somit zu der Gruppe der Goodness-of-fit-Indizes. Der TLI wird nur gering durch die Stichprobengröße beeinflusst und durch die Berücksichtigung der Freiheitsgrade werden sparsamere Modelle, bei denen weniger Parameter geschätzt werden müssen, günstiger beurteilt als weniger restriktive Modelle, bei denen mehr Parameter geschätzt werden müssen. Ein wesentlicher Nachteil des TLI ist die fehlende Normierung. Aus diesem Grund wird der TLI zwar berechnet, ist aber bei der Beurteilung der Modellpassung in den hier durchgeführten Untersuchungen zweitrangig.

Den Nachteil der fehlenden Normierung des TLI umgeht der *Comparative Fit Index* (CFI) von Bentler (1990). Der CFI ist auf das Intervall $[0;1]$ genormt. Höhere Werte deuten auf eine bessere Modellpassung hin, der TLI gehört somit zu der Gruppe der Goodness-of-fit-Indizes.

Damit sind alle Fitindizes, die in den hier durchgeführten Untersuchungen Verwendung finden, beschrieben. Es hat sich ein Quasistandard gebildet, welche Indizes bei konfirmatorischen Faktorenanalysen berichtet werden. Die im Rahmen dieser Arbeit erwähnten Indizes schließen sich der Auswahl von Brown (2006) und Bühner (2006) an, somit werden als Overall-Model-Fit-Index der RMSEA³² und SRMR sowie aus der Kategorie der auf Modellvergleichen basierenden Indizes der CFI und TLI genannt. Zusätzlich werden zu jedem Modell

³² Die Zuordnung des RMSEA in die Kategorie der Overall-Modell-Fit-Indizes ist nicht ganz eindeutig, so wird der RMSEA von Brown (2006) in die Kategorie der Modellsparsamkeitsindizes eingeordnet. Modellsparsamkeitsindizes beziehen die Anzahl der Parameter in die Berechnung eines Kennwertes ein, wobei Modelle mit weniger Parametern bevorzugt werden. Modellsparsamkeitsindizes dienen in erster Linie dem direkten Vergleich von alternativen Modellen und werden daher an dieser Stelle nicht weiter betrachtet.

der zuvor beschriebene χ^2 -Test und die Korrektur nach dem Bollen-Stine-Bootstrap sowie das Verhältnis χ^2/df erwähnt. In Tabelle 29 werden die jeweiligen Beurteilungskriterien der verwendeten Fitindizes beschrieben. Die Beurteilungskriterien werden nach der Zusammenfassung von Schermelleh-Engel et al. (2003) verwendet.

Tabelle 29: Übersicht über die Beurteilungskriterien der verwendeten Fitindizes

Fitindex	gute Modellpassung	Akzeptable Modellpassung
χ^2 -Wert	$0 \leq \chi^2 \leq 2df$	$2df < \chi^2 \leq 3df$
$p(\chi^2)$	$.05 < p \leq 1.00$	$.01 \leq p \leq .05$
χ^2/df	$0 \leq \chi^2/df \leq 2df$	$2 < \chi^2/df \leq 3$
RMSEA	$0 \leq RMSEA \leq .05$	$0,05 \leq RMSEA \leq .10$
Konfidenzintervall (CI) des RMSEA	nahe am RMSEA, linke Grenze bei 0	nahe am RMSEA
$p_{close\ fit}$	$.10 < p \leq 1.00$	$.05 \leq p \leq .10$
SRMR	$0 \leq SRMR \leq .05$	$.05 < SRMR \leq .10$
CFI	$.97 \leq CFI \leq 1.00$	$.90 \leq CFI < .97$
TLI	$.95 \leq TLI \leq 1.00$	$.90 \leq TLI < .95$

Schließlich ist in diesem Zusammenhang noch zu erläutern, wie die Interpretation der Modellpassung eines gegebenen CFA-Modells stattfindet. Die Interpretation des χ^2 -Tests und der einzelnen Indizes geschieht anhand der Tabelle 29. Der χ^2 -Test wird nur sehr vorsichtig zur Interpretation der Modellpassung verwendet. Grundsätzlich wird die Normalverteilung der Daten überprüft und bei Abweichungen innerhalb der von West et al. (1995) genannten Grenzen wird der p -Wert nach Bollen und Stine (1993) benutzt. Hauptsächlich wird der χ^2 -Wert im deskriptiven Verhältnis χ^2/df verwendet. Die anderen genannten deskriptiven Fitindizes werden ebenfalls gemäß den Beurteilungskriterien darauf hin aufgeteilt, ob ein akzeptabler oder guter Modellfit vorliegt. Wenn die verwendeten Indizes keine eindeutige Interpretation zulassen, muss eine subjektive Entscheidung getroffen werden, bei welcher weitere Faktoren wie die im nächsten Kapitel erläuterten Modifikationsindizes und Residuen, die Faktorladungen etc. berücksichtigt werden.

3.2.3 Modifikationen des Modells

Geklärt werden muss noch die Frage, was geschieht, wenn ein CFA-Modell nicht zu den empirischen Daten passt. Die Methodologie der CFA (und allgemeiner der Strukturgleichungsmodelle) ist grundsätzlich hypothesentestender Natur. Jöreskog (1993) unterscheidet drei Herangehensweisen bei der Testung von Strukturgleichungsmodellen:

- den strikt konfirmatorischen Ansatz, der davon ausgeht, dass es ein einziges Modell gibt, das getestet und dabei Modellpassung aufweist oder nicht,
- den modellvergleichenden Ansatz, bei dem mehrere alternative Modelle gegeneinander an einem einzigen Datensatz getestet werden, und den
- modellgenerierenden Ansatz, bei dem ein tentatives Modell an einem Datensatz getestet und bei fehlender Passung revidiert und erneut getestet wird, wobei die Revision des Modells entweder auf theoretischen Überlegungen oder auf Grundlage der Daten erfolgen kann.

Auf Basis des modellgenerierenden Ansatzes kann man die konfirmatorische Faktorenanalyse aber auch als ein Instrument ansehen, das zur Konstruktion und Evaluation von Skalen verwendet werden kann. Dabei bietet die CFA Vorteile gegenüber der Verwendung herkömmlicher Methoden wie Trennschärfeanalysen und explorativen Faktorenanalysen. So können zum einen mehr Parameter geschätzt werden, wie z. B. die Fehlerkomponente einer manifesten Variablen (bzw. eines Items, das zur Messung eines Konstrukts verwendet wird) und die möglicherweise vorhandene Korrelation der Messfehler. Zum anderen bieten konfirmatorische Faktorenanalysen die Möglichkeit, die Dimensionalität und Äquivalenz der Messmodelle (s. u.) zu testen. Und schließlich können mithilfe konfirmatorischer Faktorenanalysen die Reliabilitäten bestimmt werden, wenn die Voraussetzungen zum Berechnen des üblicherweise verwendeten Cronbachs α nicht gegeben sind.

Die Sichtweise der CFA als Instrument der Skalenkonstruktion ist für die Analyse der Skalen zur Messung epistemologischer Überzeugungen wichtig. Da es sich bei den Skalen hauptsächlich um eher tentative Modelle handelt (vgl. Pintrich, 2002), bei welchen in den meisten Fällen die theoretische Abgrenzung der Konstrukte nicht scharf gezogen ist, ist also nicht zu erwarten, dass die initialen CFA-Modelle sofort zu den Daten passen werden (konfirmatorischer Ansatz). Es bietet sich damit an, den modellgenerierenden Ansatz zu nutzen. Was dazu benötigt wird, sind Anhaltspunkte, nach der ein Modell modifiziert werden kann.

Ist erster Linie sollten es theoretische Gesichtspunkte sein, welche die Modifikation eines Modells bedingen. Nicht immer sind aber alle relevanten Aspekte theoretisch bedacht oder es spielen Effekte eine Rolle, die eher auf die angewandten Methoden zurückzuführen sind, wie z. B. durch positive oder negative Itemformulierung erzeugte Methodenfaktoren (DiStefano & Motl, 2009). Es sind also Herangehensweisen gefragt, die eine Untersuchung schlecht passender Modelle gestattet.

Erste Anhaltspunkte, welche der manifesten Variablen in einem Modell problematisch sind, liefern die Residuen. Dabei handelt es sich um die Differenz der empirischen Kovarianzmatrix und der modellimplizierten Kovarianzmatrix. Wird eine Kovarianz zwischen zwei manifesten Variablen durch das Modell sehr gut reproduziert, so ist diese Differenz klein. Wird eine Kovarianz durch das Modell nicht gut reproduziert, wird diese Differenz groß. Der Absolutbetrag der Residuen liefert somit einen Hinweis, ob das theoretische Modell in der Lage ist, die empirischen Zusammenhänge hinreichend genau zu beschreiben. Ein positives Residuum bedeutet, dass die empirische Kovariation unterschätzt wird und ein negatives Residuum deutet auf eine Überschätzung der empirischen Kovariation durch die modellimplizierte Kovariation hin. Allerdings sind die Residuen abhängig von der Skalierung der jeweiligen Variablen, sodass deren absolute Größe keinen Hinweis liefert, ob die Abweichung gravierend oder nur marginal ist. Eine Möglichkeit dies zu beseitigen, ist die Standardisierung der Residuen, in diesem Fall spricht man von standardisierten Residuen. Um diese zu berechnen, werden die Residuen durch ihre geschätzten Standardfehler geteilt (Brown, 2006). Die standardisierten Residuen sind dann analog zu einem z-Wert zu interpretieren. Residuen, die größer 1.96 sind, entsprechen also einem auf dem 5%-Niveau signifikanten z-Wert. Nun führt eine große Stichprobe aber im Allgemeinen zu höheren Residuen (Brown, 2006), sodass vorgeschlagen wurde, zur Untersuchung der Residuen einen größeren Wert als 1.96 anzusetzen. Byrne (1998) schlägt den Wert 2.58 vor, was einem Signifikanzniveau von 1 % entspricht. Wegen dieser Abhängigkeit der standardisierten Residuen von der Größe der Stichprobe gibt also der Absolutbetrag keinen eindeutigen Hinweis auf eine Modellfehlspezifikation und deswegen empfiehlt Brown (2006), nach Ausreißern in der Residualmatrix zu suchen. Die Inspektion der Residuen kann also Hinweise liefern, für welche Variablen im Modell Fehlspezifikationen vorliegen. Allerdings kann den Residuen nicht entnommen werden, welche Art von Fehlspezifikation vorliegt.

Genauere Hinweise auf Modellfehlspezifikationen liefern die Modifikationsindizes (Sörbom, 1989). Ein Modifikationsindex kann für jeden Parameter berechnet werden, der im Modell fixiert bzw. restringiert wurde.³³ Der Modifikationsindex ist ein Schätzwert, der angibt, um wie viel sich der χ^2 -Wert verringern würde, wenn der entsprechende Parameter frei geschätzt werden würde. Dazu wird berechnet, welchen numerischen Wert der freigesetzte Parameter

³³ Fixierte Parameter sind Parameter, die auf einen bestimmten Wert, z. B. auf 0, festgelegt wurden. Eine Art fixierter Parameter ist die auf 0 gesetzte Korrelation der Fehlervariablen im CFA-Modell in Abbildung 1. (Diese Korrelation wurde auf 0 fixiert, indem sie nicht ins Modell eingebracht wurde, es wäre ebenso möglich gewesen, sie explizit mit dem Wert 0 zu belegen.) Restringierte Parameter sind mehrere Parameter eines Modells, die ausgehend von theoretischen Überlegungen auf den gleichen numerischen Wert gesetzt wurden.

hätte. Bei der Berechnung des Modifikationsindex eines Parameteres wird ein ursprünglich im Modell fixierter oder restringierter Parameter freigesetzt, d. h. die Freiheitsgrade des Modells verringern sich um einen Parameter. Folglich wird ein Modifikationsindex dann als bedeutsam betrachtet, wenn er größer ist, als ein χ^2 -Wert bei einem Freiheitsgrad und einem Signifikanzniveau von 5 % wäre. Somit wird ein Modifikationsindex, der größer als 3.84 ist, als bedeutsam erachtet.

Ein Modifikationsindex liefert also Hinweise darauf, wie sich der χ^2 -Wert eines Modells ändern würde, wenn der restringierte oder fixierte Parameter freigesetzt werden würde. Allerdings ist der Modifikationsindex auch abhängig von der Stichprobengröße. Bei großen Stichproben führen selbst trivial kleine Werte der freigesetzten Parameter häufig zu einem numerisch hohen Modifikationsindex. Um dieses Problem zu umgehen, muss zu jedem Modifikationsindex der geschätzte Wert des Parameters betrachtet werden (Brown, 2006). Ein Modifikationsindex sollte nur dann zur Entscheidung für eine Änderung des Modells herangezogen werden, wenn die Schätzung des freigesetzten Parameters von substantzieller Größe ist.

Bei einer rein auf den Gesichtspunkt der Verbesserung der Modellpassung ausgerichteten Modifikation besteht die Gefahr von *specification searches*. Dabei wird ein Modell so lange an die Daten angepasst, bis ein guter Modellfit erreicht ist. Nach MacCallum (1986) führt ein Großteil der Modifikationen eines falsch spezifizierten Modells nicht zu der richtigen Spezifikation des Modells. Nach Silva und MacCallum (1988) führen Modifikationen, die aufgrund theoretischer Überlegungen gerechtfertigt werden können, zu besseren Ergebnissen. Aus diesem Grund empfiehlt Jöreskog (1993), nicht zwangsweise bei dem größten Modifikationsindex eine Respezifikation des Modells zu beginnen, sondern diese nur dann durchzuführen, wenn diese Modifikationen theoretisch gerechtfertigt werden können. Die Rechtfertigung kann aufgrund der substanziellen Theorie oder aufgrund einer Hilfstheorie, wie z. B. durch psychometrische Annahmen, begründet werden. Kann die durch den größten Modifikationsindex durchzuführende Änderung nicht theoretisch gerechtfertigt werden, wird beim zweit-, drittgrößten usw. Modifikationsindex weitergemacht. Wenn die Modifikation eines Modells erforderlich ist, wird jeweils immer nur eine Variable modifiziert, da die Schätzung eines Modifikationsindex immer auf der Konstanthaltung aller übrigen Modellparameter beruht. Eventuell erforderliche Modifikationen werden also schrittweise vorgenommen (Brown, 2006), d. h. ist ein Modifikationsindex vorhanden, der zusammen mit einer inhaltlichen Rechtfertigung eine Änderung des Modells nahelegt, so wird jeweils nur dieser eine Parame-

ter verändert. Weitere eventuell erforderliche Modifikationen werden auf diese Weise sukzessive vorgenommen.

Bei der Auswertung der Daten der vorzustellenden empirischen Untersuchungen werden bei einer fehlenden Modellpassung die Residuen und die Modifikationsindizes untersucht. Deuten die Modifikationsindizes auf eine Änderung des Modells hin, so wird diese nur dann vorgenommen, wenn sie inhaltlich zu rechtfertigen ist oder mit Rückgriff auf die psychometrische Theorie gerechtfertigt werden kann. Da in erster Linie CFA-Modelle untersucht werden, kann z. B. ein Item, das aufgrund seines Inhaltes nicht zu dem latenten Konstrukt passt, entfernt werden. Oder es können Items entfernt werden, welche aufgrund ihrer Formulierungen einen korrelierenden Messfehler verursachen würden.

An dieser Stelle muss noch einmal verdeutlicht werden, in welcher Weise nachfolgend konfirmatorisch vorgegangen wird. CFA-Modelle testen die Hypothese, dass eine latente Variable die Ausprägung einer oder mehrerer manifester Variablen beeinflusst. Wird nun ein Modell revidiert, so lautet die Hypothese weiterhin, dass die latente Variable die manifesten beeinflusst. Insofern wird hier immer noch hypothesentestend vorgegangen, auch wenn das Modell revidiert wird. Konfirmatorisch ist das gewählte Vorgehen in dem Sinn, dass der Einfluss einer latenten Variablen auf die manifesten Variablen geprüft werden soll. Modellgenerierend ist der Ansatz insofern, als ein a priori spezifiziertes Messmodell revidiert wird, wenn das Modell keine Passung aufweist.

Allerdings sollte immer bewusst sein, dass selbst bei der Strategie, nur Änderungen durchzuführen, die theoretisch begründbar sind, nach wie vor die Gefahr besteht, das Modell an die Besonderheiten des gegebenen Datensatzes anzupassen (MacCallum, Roznowski & Necowitz, 1992). In diesem Fall weist das Modell einen guten bis sehr guten Fit, aber auch dieser beruht nicht auf einer Stimmigkeit des Modells auf Populationsebene, sondern nur auf der gesuchten Passung des Modells in der vorliegenden Stichprobe. Das Modell beschreibt also nicht den eigentlich interessierenden theoretischen Zusammenhang in der Population, sondern lediglich die Besonderheiten der Stichprobe. Aus diesem Grund sollte bei der Verwendung von Modellmodifikationen immer eine Kreuzvalidierung erfolgen.

3.2.4 Kreuzvalidierung der CFA-Modelle

Unter Kreuzvalidierung versteht man die Überprüfung von Ergebnissen, die in einer Stichprobe anhand einer weiteren Stichprobe gewonnen wurden. Bei der ursprünglichen Kreuzvalidierung werden die Ergebnisse der ersten Stichprobe an der zweiten Stichprobe und umgekehrt die Ergebnisse der zweiten Stichprobe an der ersten Stichprobe überprüft (Bortz, 1999). Sollten keine zwei Stichproben zur Verfügung stehen, wird in Anlehnung an Brown (2006) vorgegangen, der eine ausreichend große Stichprobe in eine Analysestichprobe und eine Validierungsstichprobe aufteilt. Für die Analyse- und Validierungsstichprobe gelten die Empfehlungen zur Stichprobengröße.

Ist keine ausreichend große Stichprobe zur Aufteilung in eine Analyse- und Validierungsstichprobe vorhanden, wird auf ein eigenständiges Vorgehen zur Kreuzvalidierung zurückgegriffen. Hierbei wird zur ersten Analyse die Gesamtstichprobe verwendet. Die Gewinnung der Validierungsstichprobe erfolgte durch Aufteilen der Gesamtstichproben. Es werden dann jeweils zehn Teilstichproben per Zufall aus der Gesamtstichprobe gezogen, wobei die Teilstichproben jeweils ungefähr 75 % der Gesamtstichprobe umfassen sollten. Das Ziel dieser Prozedur ist, eine Validierung zu ermöglichen, die es gestattet, von den speziellen Eigenheiten der Analysestichprobe abzusehen und das getestete und ggf. modifizierte Modell an einer unterschiedlichen Stichprobe zu testen. Um die Möglichkeit zu minimieren, dass keine Besonderheiten der Analysestichproben in der Validierungsstichprobe zu finden sind, wird eine Auswahlquote mit 75 % gewählt, was auch dazu dienen soll, eine ausreichend große Validierungsstichprobe zu erhalten. Da aber bei einer Auswahlquote von 75 % eine große Chance besteht, auffällige Fälle per Zufall auszuwählen, wurde die Ziehung jeweils zehnmal wiederholt.

An diesen Validierungsstichproben wird dann das jeweilige CFA-Modell getestet. Die endgültige Entscheidung, ob ein Modell als validiert betrachtet werden kann, ist aber auch hier eine subjektive Entscheidung, die nicht nur danach getroffen werden kann, ob das fragliche Modell in jeder der Validierungsstichproben eine gute Anpassung aufweist. Beim Vorliegen einer ausreichend großen, unabhängigen Stichprobe wird eine herkömmliche Kreuzvalidierung durchgeführt, d. h. das in der Analysestichprobe ermittelte Messmodell wird in der Validierungsstichprobe auf seine Modellpassung hin untersucht. Ggf. können beide Methoden auch kombiniert werden.

3.3 Zusammenfassung der Auswertungsstrategien faktorenanalytischer Methoden

Im Folgenden wird die Auswertungsstrategie beschrieben, die bei der Beantwortung der Forschungsfragen zum Einsatz kommen wird. Die konfirmatorische Faktorenanalyse findet Anwendung bei den Forschungsfragen, die auf die Prüfung eines Modells abzielen. Diese Art der Fragestellung trifft z. B. auf die Prüfung der von Schommer (1990) postulierten Struktur des EBQ sowie auf die Prüfung der Faktorenstruktur des EPI zu. Bei der Durchführung einer konfirmatorischen Faktorenanalyse werden zuerst die Daten auf die Erfüllung der Voraussetzungen geprüft, was insbesondere auf die Prüfung der Verteilungsform abzielt. Berichtet wird über Abweichungen, die bei der Interpretation der Anpassungsmaße berücksichtigt werden sollen. Die Beurteilung der Modellpassung erfolgt anhand der zuvor dargestellten Fitindizes, wobei als Kriterien die Werte aus Tabelle 29 herangezogen werden. Um ein Modell als passend anzunehmen, sollen in der Regel alle Fitindizes mindestens Werte aufweisen, die eine akzeptable Modellpassung nahelegen. Allerdings können auch Modelle als passend angenommen werden, wenn ein oder mehrere Fitindizes die Werte einer akzeptablen Modellpassung unter- bzw. überschreiten, solange dafür eine plausible Erklärung gegeben werden kann.

Bei der konfirmatorischen Faktorenanalyse sollen in erster Linie inhaltliche Aspekte bei der Beurteilung der Modellpassung eine Rolle spielen. Dies wird insbesondere dann wichtig, wenn die konfirmatorische Faktorenanalyse modellgenerierend verwendet wird. Dies ist z. B. dann der Fall, wenn die konfirmatorische Faktorenanalyse als ein Mittel zur Konstruktion von Skalen verwendet wird. Bei diesem Vorgehen wird das ursprüngliche Messmodell je nach Vorliegen entsprechender Modifikationsindizes verändert, und die Modellpassung wird anschließend erneut beurteilt. Eine Modifikation soll aber nur dann regelmäßig erfolgen, wenn dies inhaltlich bzw. theoretisch zu rechtfertigen ist. Ein solches Vorgehen schließt also die Inspektion der Items ein. Ist eine Entscheidung aus inhaltlichen Gesichtspunkten nicht möglich (z. B. bei Items mit gleichem Inhalt, aber unterschiedlicher Formulierung), wird die Entscheidung aufgrund statistischer Kennwerte, z. B. der Faktorladungen der betreffenden Items, getroffen. Da auch ein solches Vorgehen interpretativ ist, wird auch dieses genau dokumentiert, um die Nachvollziehbarkeit zu gewährleisten.

Neben der konfirmatorischen Faktorenanalyse wird aber auch die explorative Faktorenanalyse eingesetzt. Die explorative Faktorenanalyse findet Anwendung, wenn die faktorielle Struktur eines Fragebogens noch unklar und nicht genau definiert ist. Forschungsfragen, die mithilfe

der EFA bearbeitet werden, sind z. B. die Prüfung der faktoriellen Struktur des EBQ auf Ebene der einzelnen Items. Die EFA kommt aber auch dann zum Einsatz, wenn ein Modell nicht mithilfe der konfirmatorischen Faktorenanalyse bestätigt werden kann, in diesem Fall kann die faktorielle Struktur mithilfe einer EFA untersucht werden. Bei der Durchführung einer EFA werden im ersten Schritt die Daten auf die Eignung zur Durchführung einer Faktorenanalyse geprüft. Dies geschieht mithilfe des KMO- bzw. MSA-Maßes. Je nach Eignung der Daten erfolgt ggf. die Untersuchung der Kommunalitäten der Variablen bzw. die Klärung der Frage, ob die Stichprobengröße überhaupt für die Durchführung einer EFA geeignet ist. Wenn erforderlich, werden die Daten soweit aufbereitet, dass eine EFA berechnet werden kann. Um die Anzahl der zu extrahierenden Faktoren zu bestimmen, wird der Scree-Plot herangezogen und ggf. eine Parallelanalyse berechnet. Bei der eigentlichen Faktorenanalyse werden die Rotationsmethode und die Faktorladungen sowie bei obliquen Rotation die Faktorkorrelationen berichtet.

Da die explorative Faktorenanalyse kein exaktes Verfahren ist und viele Schritte in der Anwendung einer EFA ein interpretatives Vorgehen verlangen und somit nicht frei von subjektiven Entscheidungen sind, wird das bei einer EFA angewandte Vorgehen so detailliert wie möglich beschrieben, um die Nachvollziehbarkeit zu gewährleisten. Überdies stehen bei der Anwendung einer EFA immer die inhaltlichen Aspekte im Vordergrund, sodass nicht immer auf die Lösung zurückgegriffen wird, die durch die beschriebenen Kennzahlen nahegelegt wird.

Die EFA findet aber auch dann Anwendung, wenn bei der Beantwortung einer Forschungsfrage die CFA zur Ablehnung eines Modells führt. In diesem Fall wird die faktorielle Struktur mithilfe der EFA ermittelt, um neben der Ablehnung des Modells weitere Aussagen über die Struktur des betrachteten Messmodells machen zu können. Die Strategie zur Auswertung ist somit zweistufig: Wenn ein Modell als nicht passend abgelehnt wird, wird daran anschließend die Struktur auf explorativem Weg ermittelt. Stehen danach ausreichend Daten zur Prüfung des auf explorativem Weg entwickelten Modells zur Verfügung, so wird in einem zweiten Schritt dieses Modell wieder mithilfe der konfirmatorischen Faktorenanalyse auf seine Modellpassung hin untersucht.

3.4 Psychometrische Eigenschaften

Eine der Fragen, welche die hier durchgeführten Untersuchungen beantworten sollen, ist die Frage nach den psychometrischen Eigenschaften der Skalen zur Messung epistemologischer Überzeugungen. Im Allgemeinen versteht man darunter die Untersuchung der Reliabilität einer Skala. In diesem Kapitel wird der aus der klassischen Testtheorie stammende Begriff der Reliabilität definiert und seine Verwendung im Rahmen der hier durchgeführten konfirmatorischen Faktorenanalysen beschrieben.

3.4.1 Definition der Reliabilität

Der Begriff der Reliabilität bezieht sich auf die Präzision bzw. Genauigkeit einer Messung. Ein Testverfahren ist perfekt reliabel, wenn die damit erhaltenen Testwerte frei von zufälligen Messfehlern sind. Das Testverfahren ist umso weniger reliabel, je größer die Einflüsse von zufälligen Messfehlern sind (Schermelleh-Engel & Wagner, 2008).

Der Zusammenhang zwischen Messwert und Messfehler wird durch die Axiome der Klassischen Testtheorie (KTT) hergestellt. Die KTT beinhaltet fünf Axiome:

- Ein beobachteter Wert (Messwert) setzt sich additiv aus dem wahren Wert und dem Fehlerwert zusammen.
- Der Erwartungswert, der Mittelwert und die Summe der Fehler sind gleich 0.
- Fehlerwert und wahrer Wert sind unkorreliert.
- Wahrer Wert und Fehlerwert zweier verschiedener Tests A und B sind unkorreliert.
- Die Fehlerwerte zweier unterschiedlicher Tests A und B sind unkorreliert.

Aus dem ersten und zweiten Axiom folgt, dass der Mittelwert der beobachteten Werte dem Mittelwert der wahren Werte entspricht. Daraus ergibt sich aber auch die Stabilität der wahren Werte, denn der einzige Wert, der schwanken kann, ist der Messfehlerwert. Insbesondere hat dies zur Konsequenz, dass sich die KTT nur dann sinnvoll anwenden lässt, wenn die wahren Werte stabil in der Zeit sind. Die KTT eignet sich somit nicht zur Erfassung von Prozessen oder Verhaltensänderungen, da die KTT von stabilen Eigenschaften ausgeht.

In der klassischen Testtheorie wird die Reliabilität nach Lord und Novick (1968) als der Quotient der Varianz der wahren Werte und der Varianz der beobachteten Werte berechnet. Allerdings ist mithilfe dieser Definition noch keine Berechnung einer Maßzahl für die Reliabilität möglich, da die Varianz der wahren Werte regelmäßig unbekannt ist. Deswegen ist es erforderlich, Methoden zur Reliabilitätsberechnung anzuwenden, dies ist üblicherweise die Be-

rechnung der internen Konsistenz nach Cronbach. Der dazugehörige Kennwert ist als Cronbachs α bekannt.

Weiterhin stellt ein berechneter Reliabilitätskoeffizient immer eine auf die jeweils erhobene Stichprobe bezogene Größe dar (Meyer, 2004). Das heißt, dass die für eine Skala ermittelte Reliabilitätskennzahl nicht als absolute Größe zu sehen ist, sondern immer eine auf die jeweils zugrunde liegende Stichprobe bezogene Maßzahl darstellt. Reliabilitätskoeffizienten können somit nicht verallgemeinert werden. Die hier berechneten Maße beschreiben daher nur die jeweilige Stichprobe.

3.4.2 Voraussetzungen der Reliabilitätsschätzung

Reliabilitätschätzungen im Sinn der internen Konsistenz nach Cronbachs α hängen von bestimmten Voraussetzungen ab. Diese müssen erfüllt sein, damit Cronbachs α als Reliabilitätskennwert sinnvoll interpretiert werden kann. Diese Voraussetzungen sind die Äquivalenz und die Eindimensionalität (bzw. Homogenität) der Messungen (Bühner, 2006, Kline, 1998).

Die Äquivalenz, d. h. Gleichwertigkeit von Messungen, wird durch die wahren Werte und die Messvarianz (z. B. Fehlervarianz) bestimmt. Deswegen unterscheidet man parallele, τ -äquivalente und kongenerische Messmodelle.

Bei kongenerischen Messungen werden die manifesten Variablen in unterschiedlichem Ausmaß durch die latente Variable beeinflusst, ebenso sind die Messfehler für jede manifeste Variable unterschiedlich. Das kongenerische Messmodell ist mit dem grundlegenden CFA-Modell aus Abbildung 12 identisch, das heißt zur Überprüfung, ob ein kongenerisches Messmodell vorliegt, kann das Modell aus Abbildung 12 direkt geprüft werden, es müssen keine Restriktionen im Modell getroffen werden. Jedes CFA-Modell, das eine Passung des theoretischen Modells an die empirischen Daten zeigt, ist somit kongenerisch.

Für τ -äquivalente Messungen gilt, dass jede der manifesten Variablen in gleichem Ausmaß von der latenten Variablen beeinflusst wird. Wie schon eine kongenerische Messung ist das τ -äquivalente Messmodell durch das grundlegende CFA-Modell aus Abbildung 12 darzustellen, wobei zur Prüfung, ob ein τ -äquivalentes Messmodell vorliegt oder nicht, die Parameter λ_i gleichgesetzt werden. Das heißt, es wird ein CFA-Modell geprüft, bei dem für das Beispiel

aus Abbildung 12 die Restriktion $\lambda_1 = \lambda_2 = \lambda_3$ gilt. Weist ein entsprechend restringiertes Modell eine ausreichende Modellpassung auf, so kann von τ -Äquivalenz ausgegangen werden.

Eine parallele Messung bedeutet, dass jede der manifesten Variablen im gleichen Ausmaß von der latenten Variablen beeinflusst wird und dass die Präzision der Messung bei jeder manifesten Variablen gleich ist. Parallelität ist übrigens auch die Bedingung, welche die Items bei der Konstruktion von Paralleltests erfüllen müssen. Zur Überprüfung, ob eine parallele Messung vorliegt, müssen im CFA-Modell einerseits die Parameter λ_i und andererseits die Messfehler δ_i gleichgesetzt werden. Das heißt, es wird ein CFA-Modell geprüft, bei dem für das Beispiel aus Abbildung 12 die Restriktionen $\lambda_1 = \lambda_2 = \lambda_3$ und $\delta_1 = \delta_2 = \delta_3$ gelten. Weist das Modell eine ausreichende Modellpassung auf, so kann von Parallelität ausgegangen werden.

Um festzustellen, ob eine kongenerische, τ -äquivalente oder parallele Messung vorliegt, müssen also CFA-Modelle mit den jeweiligen Restriktionen geprüft werden. Dabei gilt, dass ein paralleles Messmodell auch immer τ -äquivalent und kongenerisch und ein τ -äquivalentes Messmodell auch immer kongenerisch ist. Zur Prüfung würde also eine Prüfung ausreichen, ob ein Messmodell parallel ist. Allerdings kann bei einer negativen Prüfung nicht gesagt werden, ob das Messmodell dann die Bedingungen der τ -Äquivalenz erfüllt. Die Prüfung muss also in umgekehrter Richtung erfolgen. Hierzu wird in Anlehnung an ein von Brown (2006) vorgeschlagenes Verfahren vorgegangen. Danach wird bei der Passung eines CFA-Modells, das damit von selbst die Bedingung für eine kongenerische Messung erfüllt, geprüft, ob ein τ -äquivalentes Messmodell vorliegt. Dazu werden die Parameter in der o. g. Weise restringiert und das Modell wird erneut geprüft. Lassen die Fitindizes den Schluss zu, dass ein τ -äquivalentes Messmodell vorliegt, werden im letzten Schritt die Messfehler für die einzelnen manifesten Variablen gleichgesetzt. Danach wird der Fit dieses Modells überprüft. Legen die Fitindizes eine Modellpassung nahe, so ist das getestete Messmodell parallel. Um eine sinnvolle Berechnung der Reliabilität bzw. der internen Konsistenz als Schätzwert der Reliabilität zu erhalten, muss mindestens τ -Äquivalenz vorliegen.

Die Eindimensionalität eines Messmodells bezieht sich auf die Fähigkeit der Items, nur ein Merkmal zu messen. Dies wird durch zwei Sachverhalte charakterisiert (Kline, 1998), nämlich, dass ein Item nur einer latenten Variablen zugeordnet ist und dass die Messfehler der Items unkorreliert sind.

Der erste Punkt ist wichtig für die Interpretation von internen Konsistenzen. Die Idee hinter diesen Koeffizienten ist, dass jedes Item als separater Indikator für eine latente Variable angesehen wird. Aus dem Zusammenhang der Indikatoren kann die interne Konsistenz als ein Schätzwert für die Reliabilität berechnet werden. Dass die interne Konsistenz eine brauchbare Schätzung der Reliabilität einer Skala liefert, ist aber nur dann gegeben, wenn sich alle Items auf ein zugrunde liegendes Konstrukt beziehen.

Eindimensionalität wird dadurch garantiert, dass ein Item systematisch nur von einer latenten Variablen beeinflusst wird. Zur Feststellung von Eindimensionalität gibt es verschiedene Kriterien. Gerbing und Anderson (1988) haben dazu zwei Kriterien aufgestellt:

- Die Korrelation ρ zweier Indikatoren i und j , die das gleiche Konstrukt ξ messen, soll der Produktregel $\rho_{ij} = \rho_{i\xi}\rho_{j\xi}$ genügen, und
- die Korrelation ρ zweier Indikatoren i und p , wobei p ein Indikator eines anderen Konstrukts ξ^* ist, soll der Produktregel $\rho_{ip} = \rho_{i\xi}\rho_{\xi\xi}\rho_{i\xi^*}$ genügen.

Nach Steenkamp und van Trijp (1991) lässt sich Eindimensionalität auch in der Faktorladungsmatrix erkennen. Diese muss eine Einfachstruktur aufweisen, was zu der Regel führt, dass pro Zeile der Matrix nur ein Element von null verschieden sein soll. Findet sich in einem CFA-Modell also ein Hinweis, dass ein Item von mehr als einer latenten Variablen beeinflusst wird, so ist die Eindimensionalität verletzt (Kline, 1998). Eindimensionalität in diesem Sinne lässt sich aber nur dann nachweisen, wenn mehr als eine latente Variable in dem jeweiligen Modell enthalten ist. Ein Beispiel bietet das Modell in Abbildung 14, in dem Kreuzladungen vorliegen. Diese Items (manifeste Variable) werden von beiden im Modell dargestellten latenten Variablen beeinflusst. Hinweise auf Eindimensionalität nach dem Kriterium von Steenkamp und van Trijp (1991) würde in diesem Fall die Faktorladungsmatrix liefern, weist diese eine Einfachstruktur auf, so kann von Eindimensionalität ausgegangen werden. Allerdings besagt das Kriterium von Steenkamp und van Trijp nicht, ab welcher Höhe der Faktorladungen nicht von Eindimensionalität ausgegangen werden kann.

Liegen korrelierte Messfehler vor, ist dies ein Indiz, dass die Items, deren Fehler korrelieren, zusätzlich von einer anderen latenten Variablen beeinflusst werden. Wie zuvor beschrieben sind korrelierende Messfehler analog zu Methodenfaktoren, die eine weitere latente Variable bilden, welche die Items beeinflusst.

Eindimensionalität ist ein wichtiges Kriterium, da nur bei einer eindimensionalen Messung die Erfassung einer latenten Variablen ermöglicht wird. Würde z. B. ein Methodenfaktor vorliegen, bedeutete dies, dass das Messinstrument nicht nur den Einfluss des eigentlich interessierenden Merkmals erfasst, sondern auch Aspekte erfasst werden, die durch das Messinstrument bedingt sind. Dies ist aber aus diagnostischer Sicht nicht wünschenswert. Daher wird bei der Skalenkonstruktion Wert auf die Eindimensionalität im Sinne der Abwesenheit von korrelierenden Messfehlern sowie bei den Modellen, die alle Skalen enthalten, nach Möglichkeit Wert auf die Einfachstruktur der Faktorladungsmatrix gelegt.

Ob Eindimensionalität vorliegt oder nicht, kann in Anlehnung an das von Shevlin, Miles, Davis und Walker (2000) vorgeschlagene Vorgehen ermittelt werden. Hierzu werden in einem ersten Schritt eindimensionale CFA-Modelle auf ihre Modellpassung geprüft. Dadurch kann festgestellt werden, ob korrelierende Messfehler vorliegen, sodass ggf. Modifikationen des Modells erfolgen können. In einem zweiten Schritt können alle in Frage kommenden Skalen in einem gemeinsamen CFA-Modell auf Modellpassung und die Faktorladungsmatrix auf eine Einfachstruktur untersucht werden.

Sind diese Voraussetzungen der τ -Äquivalenz und Eindimensionalität erfüllt, liefert die interne Konsistenz nach Cronbachs α einen brauchbaren Schätzwert der Reliabilität und lässt sich überdies sinnvoll interpretieren.

3.4.3 Schätzungen der internen Konsistenz

Das Schätzmodell der internen Konsistenz knüpft an die Methode der Testhalbierung an (Schermelleh-Engel & Wagner, 2008), die dahingehend erweitert wird, dass eine Skala nun nicht mehr in zwei Hälften aufgeteilt wird, sondern jedes Item als eigenständiger Test angesehen wird. Die Skala wird also in so viele Teile aufgeteilt, wie sie Items enthält. Der am meisten benutzte Kennwert zur Berechnung der internen Konsistenz ist Cronbachs α . Voraussetzung dabei ist die wesentliche τ -Äquivalenz und Eindimensionalität der Items. Ist diese Voraussetzung erfüllt, ist Cronbachs α gleich der wahren Reliabilität der Skala (Osburn, 2000). Sind die Voraussetzungen nicht erfüllt, gibt diese Maßzahl eine Schätzung der wahren Reliabilität an. Bei der Berechnung ist zu beachten, dass alle Items positive Kovarianzen bzw. Korrelationen aufweisen müssen, das heißt, dass diese ggf. umkodiert werden müssen (Bühner, 2006).

Für die Interpretation von Cronbachs α sind die Bedingungen der τ -Äquivalenz und Eindimensionalität von enormer Wichtigkeit. So konnten Raykov (1997, 2001a) und Zimmermann, Zumbo und Lalonde (1993) zeigen, dass der Wert von Cronbachs α deutlich von der wahren Reliabilität eines Tests abweicht, wenn die Voraussetzungen der τ -Äquivalenz und der Eindimensionalität verletzt sind, wobei die wahre Reliabilität eines Tests durch Cronbachs α unter- oder überschätzt werden kann. Cronbachs α überschätzt die wahre Reliabilität bei kongenerischen Skalen mit korrelierenden Messfehlern, wohingegen die wahre Reliabilität bei kongenerischen Skalen ohne korrelierende Messfehler unterschätzt wird. Eine genaue Schätzung durch Cronbachs α erfolgt nur bei eindimensionalen τ -äquivalenten Skalen. Raykov (2001b) hat ein Verfahren zur Schätzung der Reliabilität ausgearbeitet, das auch auf kongenerische Messmodelle angewandt werden kann. Der hieraus resultierende Koeffizient ist als Skalenreliabilität ρ bekannt. Die Skalenreliabilität kann allerdings nicht die Schwierigkeiten der Interpretation nicht eindimensionaler Messmodelle ausgleichen.

Dem Verfahren zur Berechnung der Skalenreliabilität ρ liegt die dargestellte Definition von Reliabilität als dem Verhältnis der Varianz der wahren Werte und der Varianz der beobachteten Werte zugrunde. Zur Schätzung der Skalenreliabilität ρ wird ein Verfahren verwendet, das von Graham (2006) vorgestellt wurde. Dazu wird in einem Messmodell wie in Abbildung 12 eine weitere latente Variable X eingeführt. Diese Variable hat inhaltlich für das Modell keine Bedeutung, sondern dient nur zu Rechenzwecken, und zwar zur Berechnung der Summe der Items; um diese Berechnung durchzuführen, werden Pfade von den Items zu der Variable X modelliert und mit einem Regressionsgewicht $\lambda_i = 1$ versehen. Aus ihr lässt sich die Varianz der beobachteten Werte errechnen. Die Varianz der wahren Werte entspricht der Varianz der latenten Variablen ξ . Das Modell ist in Abbildung 16 dargestellt. Die Skalenreliabilität ρ ergibt sich, wenn die durch das Modell implizierte Korrelation der beiden latenten Variablen X und ξ mithilfe einer Schätzung nach der OLS-Methode ermittelt wird (Graham, 2006). Die quadrierte Korrelation der latenten Variablen X und ξ liefert den Reliabilitätskoeffizienten. Die angegebene Erweiterung ändert nichts an der Passung des ursprünglichen CFA-Modells, da keine neuen Parameter geschätzt werden müssen. Die Variable X ist die Summe der Itemwerte, und die Regressionsgewichte der Items wurden auf den Wert 1 fixiert. Daher ändern sich auch die Freiheitsgrade des Modells nicht und aus diesem Grund wird bei der endogen modellierten Variablen X auch auf einen Fehlerterm verzichtet. Zu jedem der bei den eigenen Untersuchungen getesteten Messmodelle wird die Skalenreliabilität ρ und Cronbachs α als Schätzwerte für die Reliabilität bestimmt.

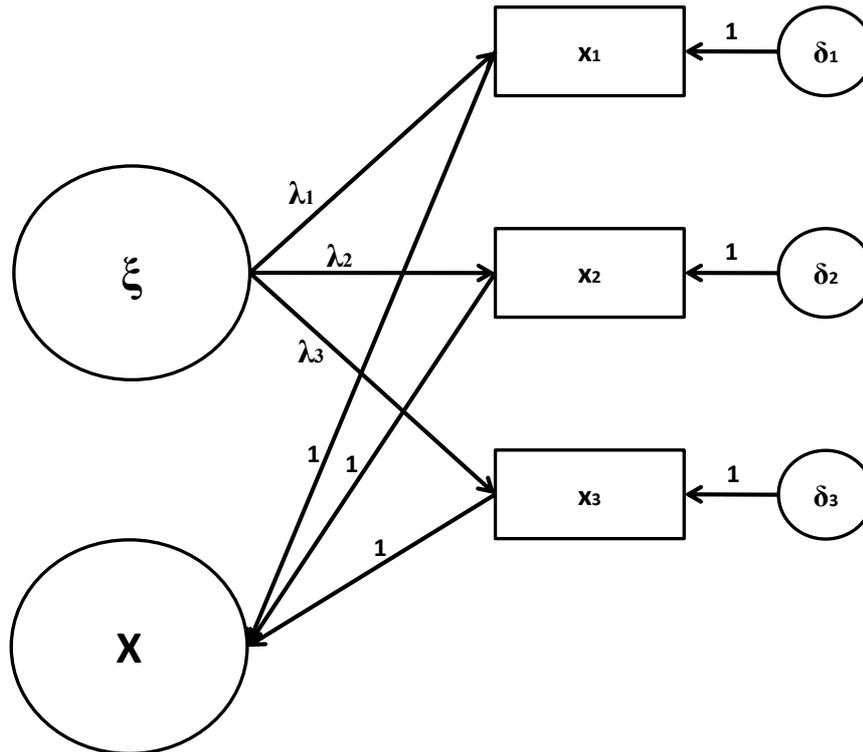


Abbildung 16: CFA-Modell zur Berechnung der Reliabilität

Es bedarf jedoch noch eines Maßstabs, um die numerische Größe von Reliabilitätskoeffizienten beurteilen zu können. Für die Beurteilung der Höhe der Reliabilität haben sich folgende Konventionen ergeben: Reliabilitäten über .90 werden als hoch bezeichnet, Reliabilitäten zwischen 0,80 und 0,90 als mittel und Reliabilitäten kleiner als .80 als niedrig (Bühner, 2006, Weise, 1975). Für die eigenen Untersuchungen wird im Anschluss an Nunnally (1978) allerdings schon eine Reliabilitätswert von .70 als ausreichend hoch betrachtet, weil es im Bereich epistemologischer Überzeugungen keine elaborierten Skalen gibt (Pintrich, 2002), die im Hinblick auf die Reliabilität mit methodisch adäquaten Mitteln optimiert wurden. So sind keine hohen Reliabilitäten bei der Untersuchung der vorliegenden Skalen zu erwarten. Außerdem sind die psychometrischen Analysen in den eigenen Untersuchungen explorativ. Das heißt, dass die hier entwickelten Skalen im Hinblick auf die psychometrischen Eigenschaften Bedarf zur Weiterentwicklung haben können und deswegen nicht vorab wegen mangelnder Reliabilität disqualifiziert werden sollten.

3.5 Auswertungsstrategie für psychometrische Fragestellungen

Bei der Prüfung der Skalen wird zuerst ein eindimensionales Messmodell auf Passung geprüft. Sollte sich keine Modellpassung zeigen, wird das Modell revidiert, wobei sich um die

Vermeidung von korrelierenden Messfehlern und damit einhergehend Methodenfaktoren bemüht wird. Durch dieses Vorgehen soll die Eindimensionalität sichergestellt werden.

Für jede entwickelte Skala wird bestimmt, welche der Äquivalenzbedingungen für diese gilt. Anschließend wird Cronbachs α als das Maß für die interne Konsistenz berechnet. Die Berechnung von Cronbachs α wird ohne Berücksichtigung der Bedingungen der τ -Äquivalenz und Eindimensionalität, welche für eine sinnvolle Interpretation notwendig sind, durchgeführt. Zusätzlich wird die Skalenreliabilität ρ berechnet, hierzu wird das zuvor beschriebene Verfahren von Graham (2006) angewandt. Dieses Vorgehen dient in erster Linie dazu, Cronbachs α mit der Skalenreliabilität ρ unter den entsprechenden Äquivalenzbedingungen vergleichen zu können.

3.6 Validität

Das Ziel dieses Unterkapitels ist es, das Konzept der Validität sowie einige zur Validierung herangezogene Methoden vorzustellen. Zuerst wird definiert, was unter Validität und Validierung zu verstehen ist und welche Implikation der Begriff der Validität hat. Anschließend werden die Methoden beschrieben, die im weiteren Verlauf eingesetzt werden sollen.

3.6.1 Definition und Bedeutung des Validitätsbegriffs

Mit Validität bezeichnet man das Ausmaß, in dem der Test die Eigenschaft misst, die er zu messen beabsichtigt (Bühner, 2006, vgl. auch Brayant, 2000). Nach Messick (1989, zitiert nach Hartig, Frey & Jude, 2008, S. 136) ist „*Validität ... ein integriertes bewertendes Urteil über das Ausmaß, in dem die Angemessenheit und die Güte von Interpretationen und Maßnahmen auf Basis von Testwerten oder anderen diagnostischen Verfahren durch empirische Belege und theoretische Argumente gestützt sind.*“. Aus dieser Sichtweise lassen sich zwei Schlussfolgerungen ableiten (vgl. Borsboom et al., 2004):

- Zum einen ist Validität ein Merkmal eines Tests. Er wird als valide betrachtet, wenn er die Erfassung einer Eigenschaft ermöglicht, die gemessen werden soll. Validität hat somit einen ontologischen Aspekt, welcher aus der Annahme hervorgeht, dass es eine Eigenschaft gibt, welche die Antwort der Probanden auf den Items kausal verursacht hat (Borsboom et al., 2004, Borsboom et al., 2009a). Diese Eigenschaft kann z. B. eine kognitive Leistung sein, welche die Antwort auf das Item eines Intelligenztests verursacht hat, oder es kann eine Persönlichkeitseigenschaft sein, welche die Einschätzung auf der Ratingskala eines Items in einem Persönlichkeitstest verursacht hat. Im Kontext der dargestellten Messmo-

delle kann auch von einer latenten Variablen gesprochen werden, welche die Antworten eines Probanden verursacht hat.

- Zum anderen wird dem Test Validität zugesprochen oder nicht zugesprochen, nachdem eine Reihe von Untersuchungen zur Überprüfung der Validität angestellt wurde. Diese Überprüfung ist die Aktivität, die man als Validierung bezeichnet. Validierung als Aktivität spiegelt einen erkenntnistheoretischen Aspekt wider. Die dabei zugrunde gelegte erkenntnistheoretische Sichtweise ist der Entitätenrealismus (Borsboom et al. 2003, vgl. Kapitel 1.7).

Wichtig festzuhalten ist, dass nach dem Entitätenrealismus die latente Variable als in realiter existierend betrachtet wird und dass zur Prüfung der Validität des Tests, der zur Messung dieser latenten Variable herangezogen wird, eine Validierung durchgeführt werden muss. Die nächsten Kapitel handeln dementsprechend von den Methoden, die im Prozess der Validierung eingesetzt werden. Validität ist allerdings keine einheitliche Eigenschaft eines Tests, sondern man unterscheidet verschiedene Arten der Validität. Eine grundsätzliche Aufteilung unterscheidet zwischen Inhaltsvalidität und Konstruktvalidität (Meyer, 2004). Inhaltsvalidität bezieht sich auf die Frage, ob die Items zur Erfassung eines Konstrukts mit der theoretischen Definition des Konstrukts übereinstimmen. Einheitliche Methoden zur Beurteilung von Inhaltsvalidität gibt es nicht (vgl. Fisseni, 2004). Üblicherweise wird Items dann Inhaltsvalidität zugesprochen, wenn Experten sich über die Zuordnung von Items zu latenten Variablen bzw. der Zuordnung von Items zu der Definition eines interessierenden Konstrukts einig sind. Konstruktvalidität bedeutet, dass sich Hypothesen empirisch überprüfen lassen, die aus der theoretischen Konzeptualisierung des interessierenden Konstrukts abgeleitet worden sind. Nachfolgend werden die verschiedenen Facetten der Konstruktvalidität sowie die zur Validierung verwendeten Methoden beschrieben.

3.6.2 Konstruktvalidität auf Itemebene

Ausgehend von der theoretischen Definition eines Konstrukts lassen sich Hypothesen über die Zusammenhänge der Items ableiten, die von dem Konstrukt kausal beeinflusst wurden. Von primärem Interesse bei der Prüfung der Konstruktvalidität ist, ob Unterschiede bei der Beantwortung der Items kausal durch Unterschiede des zugrunde liegenden Konstrukts verursacht werden (Borsboom et al., 2003, 2004, Borsboom et al., 2009a, Hartig et al., 2008). In Bezug auf die Konstruktvalidität auf Itemebene muss somit geklärt werden, ob das Antwortverhalten der Probanden auf den Items durch ein gemeinsames, zugrundeliegendes Konstrukt verur-

sacht wird. Die kausale Verursachung des Antwortverhaltens durch das latente Konstrukt ist auch gemeint, wenn dann von Validität gesprochen wird, wenn ein Test misst, was er zu messen vorgibt (Borsboom et al., 2009). Wird Validität in diesem Sinn definiert, so bedeutet dies, dass sich die Werte eines Messinstruments nur dann verändern, wenn sich das zu messende Konstrukt ändert. Zur Prüfung von Hypothesen über die Zusammenhangstruktur von Items bieten sich die in Kapitel 3.2.1 vorgestellten CFA-Modelle an. Dem Vorschlag von Shevlin et al. (2000) folgend muss geprüft werden, ob eine Skala eindimensional ist und ob korrelierte Messfehler vorliegen. Damit kann gezeigt werden, dass das Antwortverhalten auf den Items von der latenten Variablen kausal bedingt wird (Pearl, 2000).

Dabei ist es aber nicht erforderlich, dass sich die Prüfung auf ein eindimensionales CFA-Modell beschränkt. Es kann auch auf zwei- oder mehrdimensionale CFA-Modelle geprüft werden. Wichtig zur Bestätigung der Zusammenhangshypothese ist dabei nicht nur eine ausreichende Modellpassung, sondern insbesondere auch, dass die Items nur durch einen Faktor beeinflusst werden. Ein mehrdimensionales CFA-Modell liefert somit einen stärkeren Nachweis für die Konstruktvalidität eines Tests, als dies ein CFA-Modell mit nur einem Faktor leisten kann. Allerdings muss an dieser Stelle angemerkt werden, dass es sich bei dieser Art der Prüfung von Konstruktvalidität lediglich um eine notwendige Bedingung handelt (Hartig et al., 2008), ob tatsächlich Konstruktvalidität vorliegt oder nicht, kann letztendlich mithilfe dieser Methoden nicht entschieden werden.

3.6.3 Konstruktvalidität auf Ebene des Gesamttests

„*Konstruktvalidität* umfasst die empirischen Befunde und Argumente, mit denen die Zuverlässigkeit der Interpretation von Testergebnissen im Sinne erklärender Konzepte, die sowohl die Testergebnisse selbst als auch die Zusammenhänge der Testwerte mit anderen Variablen erklären, gestützt wird“ (Messick, 1995, zitiert nach Hartig et al., 2008, S. 145). Konstruktvalidität auf Ebene des Gesamttests bezieht sich somit auf die Zusammenhänge eines Konstrukts mit anderen Konstrukten. Diese Art der Konstruktvalidierung beruht auf dem nomologischen Netz (Cronbach & Meehl, 1955). Die Idee des nomologischen Netzwerks ist, dass die Bedeutung eines Konstrukts durch die Beziehung mit anderen Konstrukten erklärt wird. Dazu werden zwei verschiedene Bereiche unterschieden: zum ersten der Bereich der Theorie und zum anderen der Bereich der Beobachtungen (Hartig et al., 2008). Im Bereich der Theorie sind die theoretisch definierten Konstrukte angesiedelt, ebenso sind dort auch die Annahmen über die gesetzmäßigen Zusammenhänge eines Konstrukts mit anderen Konstrukten formu-

liert. Diese gesetzmäßigen Zusammenhänge bilden das sogenannte nomologische Netz. Die theoretischen Konstrukte sind nicht beobachtbar. Im Bereich des Beobachteten werden Variablen spezifiziert, die beobachtbar sind und durch Korrespondenzregeln mit den Konstrukten im Bereich der Theorie verbunden sind. Um nun Konstruktvalidität nachzuweisen, müssen die Beziehungen der Variablen im Bereich des Beobachteten mit den Aussagen über die Beziehungen der Konstrukte im Bereich der Theorie übereinstimmen.

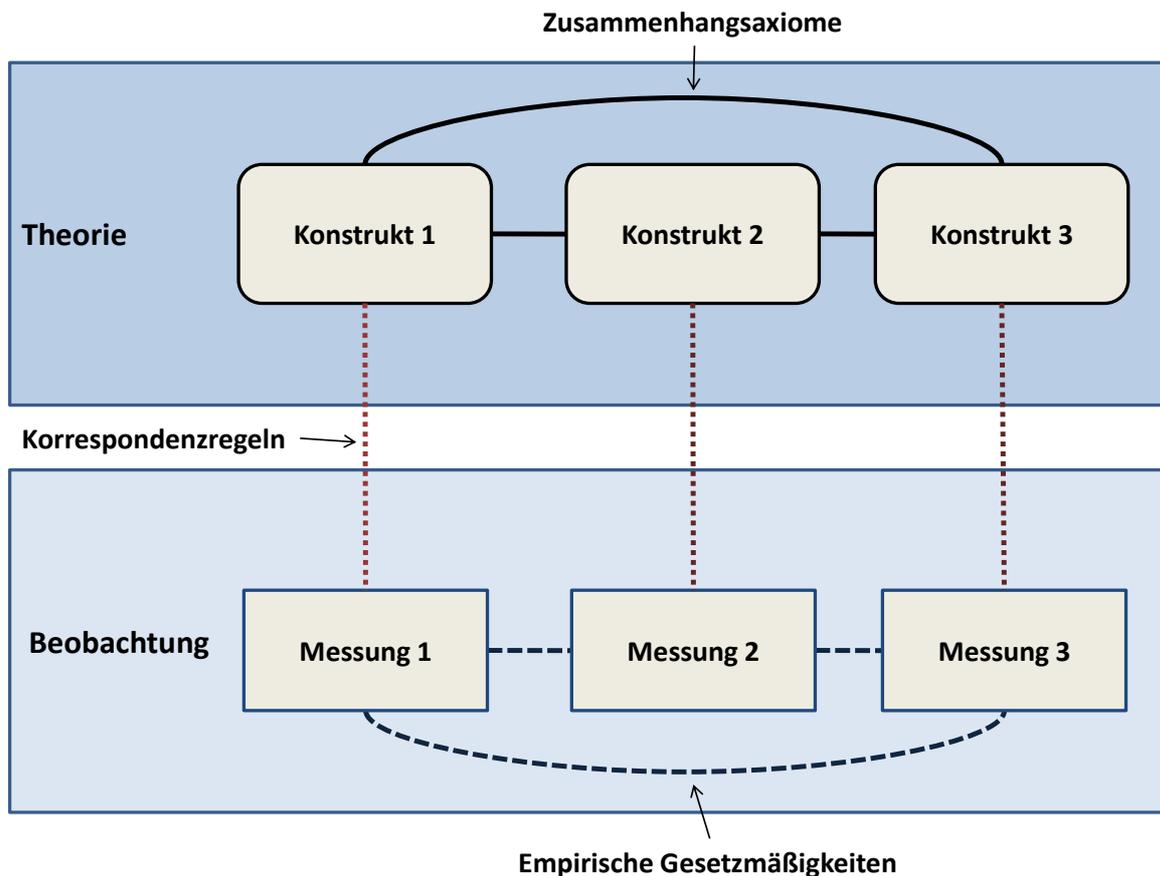


Abbildung 17: Nomologisches Netz (in Anlehnung an Hartig et al., 2008)

Abbildung 17 verdeutlicht die Idee des nomologischen Netzes. Die theoretischen Konstrukte sind durch abgerundete Rechtecke dargestellt und die Beziehung der Konstrukte durch die durchgezogenen Linien. Die Variablen im Bereich der Beobachtung sind durch Rechtecke dargestellt und die Beziehungen dieser Variablen durch gestrichelte Linien. Die gepunkteten Linien stellen die Korrespondenzregeln der Konstrukte mit den beobachteten Variablen dar.

Das nomologische Netz – und insbesondere die Idee, dass es Elemente gibt, die beobachtet werden können, und solche, die es aufgrund theoretischer Überlegungen gibt – findet eine Entsprechung in der Standardkonzeption von Theorien in den Sozialwissenschaften (Westermann, 2000). In der Standardkonzeption enthält eine Theorie Aussagen über die Zusammenhänge von theoretischen Begriffen. Ein weiteres wesentliches Element dieser Standardkon-

zeption ist das Zweisprachenkonzept. Danach gibt es eine theoretische Sprache und eine Beobachtungssprache. Die theoretische Sprache enthält Begriffe, die Aussagen der Theorie bezeichnen. Die Beobachtungssprache enthält nur Begriffe, die direkt Beobachtbares bezeichnen (Westermann, 2000). Zu diesen beiden Sprachelementen gibt es Zuordnungsregeln, welche die Begriffe der Beobachtungssprache denen der theoretischen Sprache zuordnet.

Diese Zweisprachenkonzeption liegt aber auch der Methodik der Strukturgleichungsmodelle zugrunde (Jöreskog & Sörbom, 1993). Strukturgleichungsmodelle sind eine Erweiterung der bereits dargestellten CFA-Modelle. Bei CFA-Modellen geht es um die Beziehung von latenten Variablen und den durch sie bedingten manifesten Variablen. Diese Beziehung bezeichnet man als *Messmodell*. Da aber auch Beziehungen zwischen den latenten Variablen beschrieben werden können, werden diese Beziehungen als *Strukturmodell* bezeichnet.

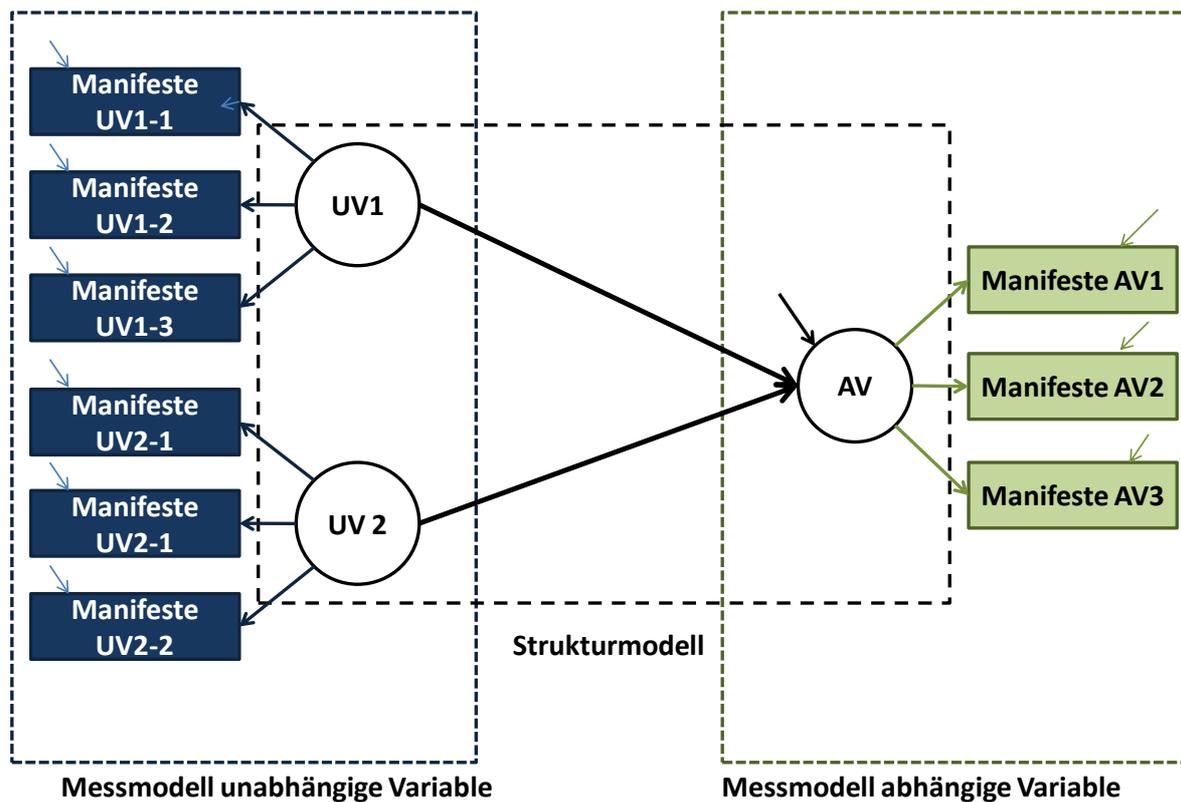


Abbildung 18: Schematische Darstellung des Messmodells und Strukturmodells

Das Messmodell leistet die Zuordnung von manifesten und latenten Variablen. Im Strukturmodell werden die Beziehungen zwischen den latenten Variablen spezifiziert, die in exogene (unabhängige) und endogene (abhängige) Variablen eingeteilt werden. Die exogenen latenten Variablen werden also zur Erklärung der endogenen latenten Variablen herangezogen. Dies

wird in Abbildung 18 durch den Pfeil dargestellt.³⁴ Den Messmodellen liegt ein dem der CFA analoger faktorenanalytischer Denkansatz zugrunde. Das heißt, in den Messmodellen werden die angenommenen Zusammenhänge der manifesten mit den latenten Variablen beschrieben. Dem Strukturmodell liegt hingegen ein regressionsanalytischer Ansatz zugrunde (Hodapp, 1994). Die hier ermittelten Parameter lassen sich wie Regressionsgewichte interpretieren, auch ist eine Signifikanzprüfung dieser Parameter möglich und wie bei den Faktorladungen des Messmodells ist ebenso eine Standardisierung möglich.

Diese im Mess- und Strukturmodell dargestellten Zusammenhänge drücken sich wie bei der CFA durch die Kovarianzmatrix der manifesten Variablen aus. Diese kann mithilfe der beschriebenen Methoden auf Modellpassung hin untersucht werden. Darüber hinaus werden Parameterschätzungen ähnlich denen der CFA abgegeben, die in Kapitel 3.2.1 genauer dargestellt werden. Ergibt das Modell eine ausreichende Passung und stimmen die Parameterschätzungen mit den theoretisch angenommenen Zusammenhängen überein, erlaubt dies den Rückschluss auf die Konstruktvalidität auf Ebene des Gesamttests. Strukturgleichungsmodelle bieten also die Grundlage zur Untersuchung der Konstruktvalidität.

Für die Konstruktvalidierung der Skalen zur Erfassung von epistemologischen Überzeugungen wird dem in Kapitel 1.6.1 vorgestellten Modell von Hofer (2001) entsprechend die Beziehung zu Lernstrategien untersucht. Nach diesem Modell beeinflussen die epistemologischen Überzeugungen die Lernstrategien und weiterhin wird angenommen, dass dieser Vorgang kausal gerichtet ist, d. h. die Wirkrichtung geht von den epistemologischen Überzeugungen in Richtung auf die Lernstrategien. Es kann somit ein Strukturmodell aufgestellt werden, welches die epistemologischen Überzeugungen als exogene und die Lernstrategien als endogene Variable enthält. Gleichzeitig lässt sich zu den latenten Variablen im Strukturmodell auch ein Messmodell formulieren, bei dem Items zur Messung von epistemologischen Überzeugungen und Lernstrategien den jeweiligen Konstrukten zugeordnet werden.

Um die Methode zu verdeutlichen und die Interpretation der geschätzten Parameter zu erläutern, wird im Folgenden ein Modell mit zwei latenten exogenen und einer latenten endogenen

³⁴ Hier wird eine gerichtete Beeinflussung der endogenen Variablen durch die exogene Variable dargestellt. Es lässt sich aber auch eine Abhängigkeit im Sinne einer Kovarianz modellieren. Der Pfeil würde dann durch einen Doppelpfeil ersetzt werden (vgl. Backhaus et al., 2003), insofern stellt schon das CFA-Modell aus Abbildung 13 ein in Mess- und Strukturmodell aufteilbares Strukturgleichungsmodell dar. Im Folgenden wird aber von gerichteten Zusammenhängen ausgegangen, so dass nur diese hier betrachtet werden. Allerdings ist es nicht allgemeingültig, bei der Konstruktvalidierung nur von gerichteten Zusammenhängen auszugehen.

Variablen dargestellt. Jeder latenten Variablen sind dabei drei manifeste Variablen zugeordnet. Das Modell ist in

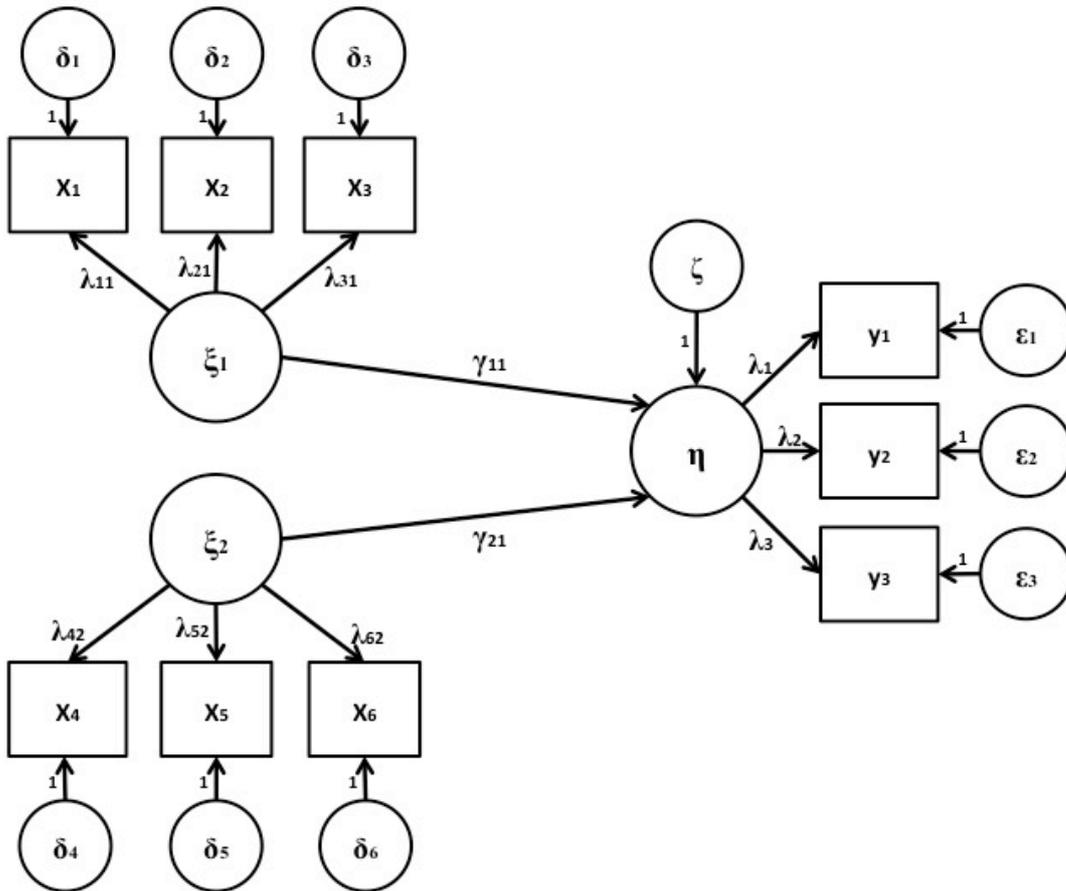


Abbildung 19 dargestellt.

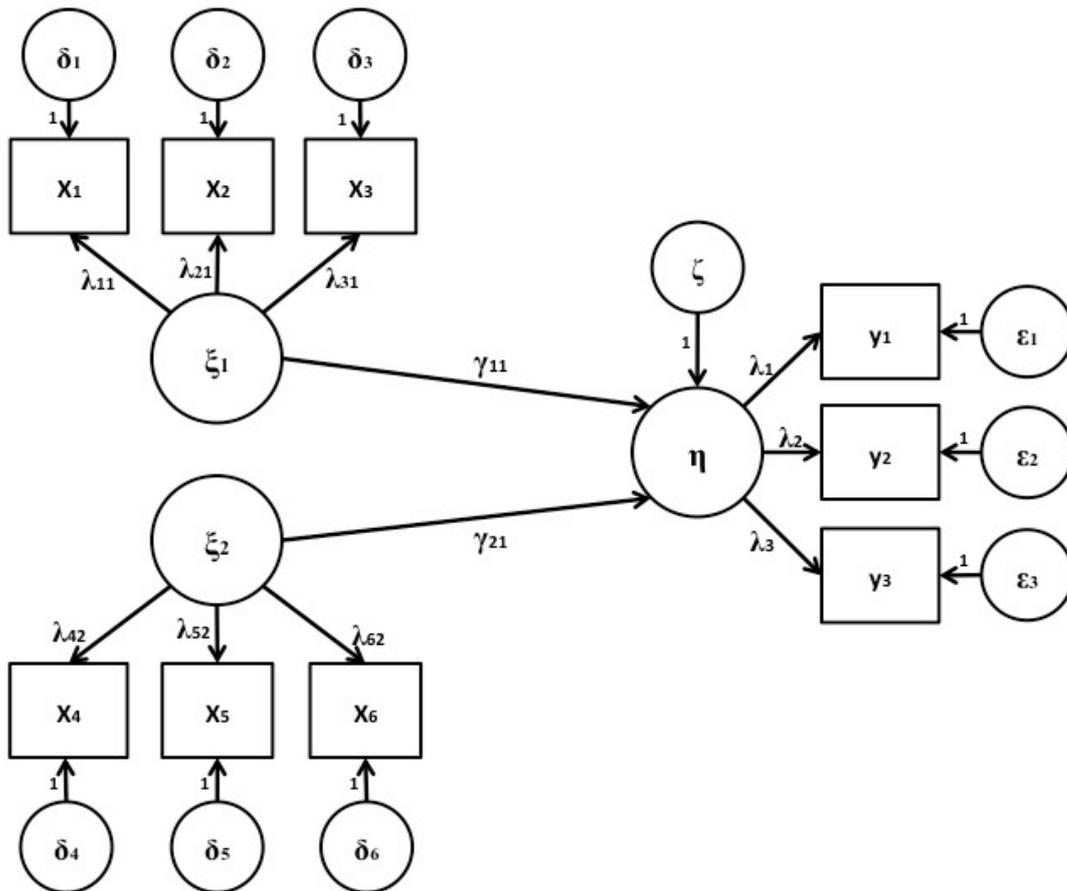


Abbildung 19: Strukturgleichungsmodell zur Konstruktvalidierung

Das abgebildete Strukturgleichungsmodell ist prototypisch für die im weiteren Verlauf verwendeten, allerdings wird aufseiten der exogenen Variablen die Anzahl nicht auf zwei beschränkt sein. Als exogene Variablen werden die epistemologischen Überzeugungen verwendet, aufseiten der endogenen Variablen wird jeweils eine Variable eingefügt, bei der es sich um eine Lernstrategie handelt. Für die Bezeichnung der Variablen in Abbildung 19 wird die übliche LISREL-Notation verwendet (vgl. Backhaus et al. 2003).

Die Parameter γ_i sind die Regressionsgewichte, die den Einfluss der exogenen auf die endogene Variable beschreiben. ζ ist die Fehlervariable, ähnlich wie bei den manifesten Variablen in einem CFA-Modell wird bei den endogenen Variablen angenommen, dass diese sich nicht perfekt durch die exogenen Variablen erklären lassen. Die Fehlervariable beschreibt diesen nicht erklärten Anteil. Die Regressionsgewichte zwischen den exo- und endogenen Variablen entsprechen den Regressionsgewichten in einer herkömmlichen linearen Regression. Sie geben den Einfluss der exogenen Variablen auf die endogene Variablen an, wobei es sich auch hier um partielle Regressionsgewichte handelt. Entsprechend können die Regressionsgewichte auch wie in einer linearen Regression interpretiert werden und die Beurteilung der Größe

eines Regressionskoeffizienten ist ebenso mithilfe der Einteilung nach Cohen (1988) möglich. Dieser Einteilung nach werden Regressionskoeffizienten von .10 bis .30 als klein bezeichnet, von .30 bis .50 als mittel und größer als .50 als groß. Die standardisierten Regressionsgewichte γ_i sowie die standardisierte Fehlervariable ζ können mittels des *critical ratio* auf Signifikanz geprüft werden.

Ebenso wie in linearen Regressionsmodellen üblich können auch die Varianzaufklärungen berechnet werden, wobei hier der Varianzanteil der endogenen Variablen berechnet wird, der sich durch die exogenen Variablen aufklären lässt. Diese Varianzaufklärung wird als *squared multiple correlation* (smc) bezeichnet und ist wie der quadrierte multiple Korrelationskoeffizient R^2 zu interpretieren. Allerdings kann nicht wie bei herkömmlichen Regressionsmodellen üblich eine Signifikanztestung mittels eines F-Tests berechnet werden. Auch auf die *squared multiple correlation* lässt sich analog einer einfacheren Regression die Klassifikation von Cohen (1988) anwenden. Zum Vergleich mit dieser Klassifikation muss allerdings die Quadratwurzel aus der *squared multiple correlation* berechnet werden, die wie der multiple Korrelationskoeffizient R betrachtet werden kann. In der Klassifikation von Cohen werden die Korrelationskoeffizienten zwischen .14 und .36 als klein, zwischen .36 und .51 als mittel und größer als .51 werden als groß bezeichnet.

Strukturgleichungsmodelle stellen eine Verallgemeinerung des Allgemeinen Linearen Modells dar (Moosbrugger, 2002). Der Vorteil gegenüber einer klassischen Regressionsanalyse liegt aber in der Kopplung von Mess- und Strukturmodell. Diese Verbindung erlaubt es, dass beim Strukturmodell die Reliabilität der Messmodelle berücksichtigt wird (wohingegen die klassische Regressionsanalyse annimmt, dass die Prädiktoren fehlerfrei erfasst werden). Die Berücksichtigung der Reliabilität der Variablen ist wichtig wegen des Verdünnungsparadox (Fisseni, 2004). Dieses besagt, dass die Korrelation von Prädiktor und Kriterium sinkt, wenn die Reliabilität von Prädiktor und Kriterium steigt. Umgekehrt steigt die Korrelation von Prädiktor und Kriterium, wenn die Reliabilität von Prädiktor und Kriterium sinkt. Um diesen Effekt auszugleichen, muss eine Minderungskorrektur durchgeführt werden (Fisseni, 2004). Die in einem Strukturgleichungsmodell berechneten Regressionsgewichte berücksichtigen diese Minderungskorrektur (DeShon, 1998).

So weit die Darstellung des Validierungsprozesses, wie er für die Prüfung der Annahmen des Zusammenhangs von epistemologischen Überzeugungen und Lernstrategien angewandt wird.

Das Modell von Hofer (2001) gestattet es, Aussagen über die kausale Beeinflussung von Lernstrategien durch epistemologische Überzeugungen zu machen. Entsprechend ist auch das nomologische Netz ausgerichtet und die Methoden zur empirischen Prüfung des nomologischen Netzes müssen die Wirkrichtung berücksichtigen, epistemologische Überzeugungen stellen in diesem Fall die exogenen und die Lernstrategien die endogenen latenten Variablen dar. Auch diese Hypothesen lassen sich mit Strukturgleichungsmodellen überprüfen. Dabei wird die gerichtete Wirkung der exogenen auf die endogene Variable durch einen Zusammenhang zwischen zwei latenten Variablen ersetzt. Im Strukturmodell muss also ein Pfad zwischen den exogenen und den endogenen latenten Variablen modelliert werden. Zusätzlich können Zusammenhänge zwischen den exogenen latenten Variablen als Kovarianz modelliert werden. In Abbildung 20 ist ein Strukturgleichungsmodell mit drei latenten Variablen und paarweisen Kovarianzen dargestellt.³⁵ Modelle wie in Abbildung 20 werden zur Konstruktvalidierung des EPI eingesetzt, wobei die Anzahl der latenten und manifesten Variablen variieren kann. Die Prüfung der Konstruktvalidität erfolgt in diesem Fall über die Fitindizes des Modells sowie den Pfadkoeffizienten und die *squared multiple correlation*.

³⁵ An dieser Stelle erfolgt eine kurze Anmerkung zu Strukturgleichungsmodellen. Die zu Beginn dargestellten CFA-Modelle sind nur eine besondere Gruppe von Strukturgleichungsmodellen. Der Schwerpunkt bei CFA-Modellen liegt auf der Prüfung der Messmodelle und ggf. der Prüfung der korrelativen Struktur der Faktoren, während bei Strukturgleichungsmodellen die Zusammenhänge auf der Ebene der latenten Variablen den Schwerpunkt bilden und auch gerichtete Beeinflussungen der latenten Variablen beinhalten. CFA-Modelle mit mehr als einer latenten Variablen lassen sich in Messmodell und Strukturmodell aufgliedern, so bilden z. B. die Beziehungen zwischen den latenten Variablen in Abbildung 13 und Abbildung 20 das Strukturmodell und die Zuordnungen der manifesten Variablen zu den latenten Variablen die Messmodelle.

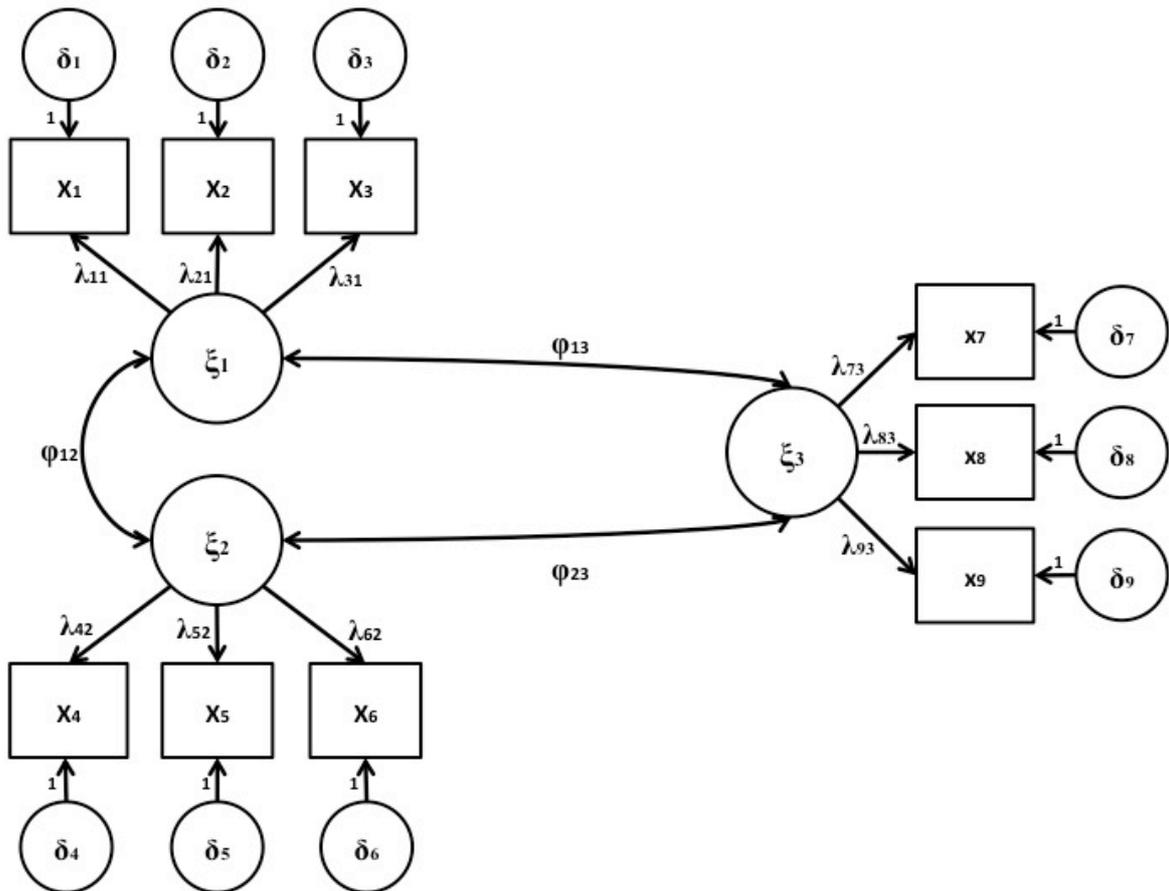


Abbildung 20: Strukturgleichungsmodell mit drei latenten Variablen und Kovarianzen

Damit ist die Darstellung der Strukturgleichungsmodelle, wie sie hier zur Konstruktvalidierung benutzt werden, abgeschlossen. An dieser Stelle muss jedoch eine Präzisierung des Validitätsbegriffs vorgenommen werden. Zuvor wurde erläutert, dass die exogenen Variablen zur Erklärung der endogenen Variablen herangezogen werden, im Sinne einer Regression stellt also die exogene Variable den Prädiktor und die endogene Variable das Kriterium vor. Wird eine Variable zur Vorhersage einer anderen Variablen benutzt, spricht man auch von Kriteriumsvalidität (vgl. Amelang & Zielinski, 2001). Die Idee ist, dass die Testwerte zur Vorhersage eines Kriteriums benutzt werden. Kriteriumsvalidität kann durch empirische Zusammenhänge mit den spezifizierten Kriterien belegt werden. Je enger die Zusammenhänge sind, desto besser kann die Kriteriumsvalidität als belegt gelten. Im Gegensatz zur Konstruktvalidierung mit der Idee des nomologischen Netzes, welches den Zusammenhang zwischen verschiedenen Konstrukten auf theoretischer Ebene spezifiziert, müssen bei der Ermittlung der Kriteriumsvalidität die Konstrukte nicht in einem theoretisch herleitbaren Zusammenhang stehen. Ist bei der Prüfung der Kriteriumsvalidität der Zusammenhang zwischen Prädiktor und Kriterium jedoch theoretisch ableitbar, so kann der gleiche empirische Zusammenhang zur Überprüfung der Kriterien- als auch der Konstruktvalidität herangezogen werden (Hartig et

al., 2008). Bei der Kriteriumsvalidität wird zwischen prädiktiver und konkurrender Validität unterschieden (Fisseni, 2004). Werden die Prädiktoren zeitlich vor dem Kriterium erhoben, so wird von prädiktiver Validität gesprochen. Werden Prädiktoren und das Kriterium zur gleichen Zeit erhoben, wird von konkurrender Validität gesprochen.

3.7 Auswertungsstrategie zur Validierung des EPI-Fragebogens

Bei der letzten Forschungsfrage zum EPI-Fragebogen behandelt die Beziehungen der EPI-Skalen zu Lernstrategien. Um die Kriteriumsvalidität der EPI-Skalen zu prüfen, wird wie bereits beschrieben mithilfe eines Regressionsmodells, bei dem die EPI-Skalen als Prädiktoren und die LIST-Skalen als Kriterien enthalten sind, der Zusammenhang zwischen epistemologischen Überzeugungen und Lernstrategien erfasst.

Bei der Auswertung sind zwei Aspekte zu berücksichtigen. Der erste Aspekt ist die Modellpassung der jeweiligen Regressionsmodelle. Um die Modellpassung zu beurteilen, werden die Fitindizes berichtet, die im Rahmen der konfirmatorischen Faktorenanalyse vorgestellt wurden. Die Beurteilung der Indizes erfolgt nach den gleichen Kriterien, die auch für die konfirmatorischen Faktorenanalyse angewandt wurden. Der zweite Aspekt neben der Modellpassung ist die Beurteilung der Stärke des Zusammenhangs zwischen den EPI-Skalen und den Lernstrategien. Um die Stärke dieses Zusammenhangs zu beurteilen, werden der multiple quadrierte Korrelationskoeffizient und die einzelnen Regressionskoeffizienten betrachtet.

Der Zusammenhang zwischen epistemologischen Überzeugungen und Lernstrategien wird aber nur dann untersucht, wenn das Regressionsmodell eine ausreichende Modellpassung aufweist. Bei der Prüfung der Kriteriumsvalidität sind also zwei Szenarien zu unterscheiden. Zum einen kann es vorkommen, dass ein Regressionsmodell keine Passung zeigt und somit das angenommene Zusammenhangsmodell nicht verwendet werden kann. In diesem Fall ist es nicht möglich, die Zusammenhänge epistemologischer Überzeugungen und Lernstrategien mithilfe eines solchen Regressionsmodells zu beschreiben. Zum anderen kann es sein, dass das Regressionsmodell eine ausreichende Passung aufweist. In diesem Fall können die Zusammenhänge untersucht werden. Dabei kann es aber vorkommen, dass keine Zusammenhänge zwischen den Prädiktoren und dem Kriterium gefunden werden. Eine Modellpassung garantiert daher nicht, dass auch signifikante bzw. ausreichend starke Zusammenhänge gefunden werden.

4. Ergebnisse zum EBQ

4.1 Faktorenanalytische Auswertung des EBQ

Dieses Kapitel befasst sich mit den Fragen, die im Hinblick auf das Modell von Schommer und den EBQ als den dazugehörigen Fragebogen formuliert wurden. Dies waren im Einzelnen die folgenden Forschungsfragen:

- *Lässt sich das Modell von Schommer durch eine konfirmatorische Faktorenanalyse empirisch bestätigen?*
- *Wie ist die faktorielle Struktur des Schommer-Fragebogens bei Faktorisierung auf der Ebene der einzelnen Items?*
- *Welche Skalen lassen sich aus den Ergebnissen der Faktorenanalyse ableiten und welche psychometrischen Eigenschaften besitzen diese Skalen?*

Diese Forschungsfragen werden im Folgenden mithilfe der vorgestellten Methoden untersucht und anschließend diskutiert.

4.1.1 Vorbereitung der Daten

Insgesamt wurde der Schommer-Fragebogen von 223 Personen bearbeitet, wobei die in Anhang A aufgeführte deutsche Version des EBQ verwendet wurde. Zusätzlich zum EBQ wurde den Probanden eine Version des EPI-Fragebogens von Moschner und Gruber (2005)³⁶ vorgelegt.

Zu Beginn des Fragebogens erfolgte eine ausführliche schriftliche Instruktion zum Thema und zum Bearbeiten des Fragebogens selbst. Die Instruktion war so gestaltet, dass eine selbstständige Bearbeitung des Fragebogens ohne Zutun eines Versuchsleiters möglich sein sollte. Die Beantwortung der Items erfolgte auf einer 6-stufigen Ratingskala mit den Stufen *stimmt überhaupt nicht* – *stimmt weitgehend nicht* – *stimmt eher nicht* – *stimmt ein wenig* – *stimmt weitgehend* – *stimmt genau*. Im Anschluss an die Bearbeitung der Skalen erfolgte die Erhebung der demografischen Daten. Abgefragt wurden das Alter, das Geschlecht, der höchste Schulabschluss, das Studienfach und die Anzahl der Semester bzw. die Berufsausbildung, die Muttersprache und das Herkunftsland, die Gesamtnote des Schulabschlusses, die letzte Mathematiknote sowie die Anzahl der Geschwister.

³⁶ Es handelt sich hierbei nicht um den EPI-Fragebogen, der nachfolgenden Kapitel 5 gesondert besprochen werden soll, sondern um eine Vorform dieses Fragebogens.

Zusätzlich zu der Printversion wurde auch eine Onlineversion gestaltet, wobei auch hier der EBQ sowie der EPI-Fragebogen von Moschner und Gruber (2005) verwendet wurde. Es wurden dieselben demografischen Merkmale abgefragt.

Der Fragebogen wurde im Sommersemester 2007 an Lehramtsstudierende und Studierende der Psychologie am Campus Saarbrücken sowie an Studierende der Medizin am Campus Homburg der Universität des Saarlandes verteilt. Weiterhin wurde der Fragebogen im Jahr 2008 an Teilnehmer von Seminaren einer Berufsgenossenschaft im Saarland verteilt. Im Wintersemester 2008 wurde der Fragebogen an Lehramtsstudierende sowie Studierende der Psychologie an der Universität Regensburg verteilt. Die Onlineversion wurde von November 2009 bis Februar 2010 unter dem Titel „Einstellungen zu Wissen und Lernen“ auf der URL <http://www.internet-psychologie.net/> dargeboten.

Die Printversion des Fragebogens wurde von insgesamt 221 Versuchspersonen bearbeitet. Davon wurden 14 Fragebogen ausgeschlossen, weil die Fragebögen unvollständig ausgefüllt worden waren bzw. die demografischen Angaben ausgelassen wurden, wobei sechs fehlende Werte bei den EBQ-Items toleriert wurden. Die Wahl des Kriteriums von sechs fehlenden Werten beruht auf der Annahme, dass das Fehlen von mehr als 10 % der Antworten auf ein willkürliches Antwortverhalten hindeuten kann und der Datensatz des betreffenden Probanden nicht für eine Auswertung geeignet ist, wobei gleichzeitig dem Probanden aber ein gewisser Anteil an ausgelassenen Antworten (z. B. durch Übersehen eines Items) zugestanden wird. Die Onlineversion wurde von 28 Versuchspersonen bearbeitet, wobei 12 Datensätze aufgrund 10% fehlender Werte bzw. spaßhafter Antworten (z. B. Beruf „Diplom-Sozialschmarotzer“) ausgeschlossen wurden. Insgesamt ergaben sich somit die Daten von 223 Versuchspersonen, welche den EBQ vollständig bearbeitet hatten. Der Mittelwert des Alters für alle Versuchspersonen betrug 27 Jahre ($SD = 8.37$, s. Tabelle 30 für eine Übersicht).

Tabelle 30: Deskriptive Statistik der Altersverteilung in Jahren

Geschlecht	<i>n</i>	Range	Min	Max	<i>M</i>	<i>SD</i>
männlich	86	38	19	57	30.06	10.00
weiblich	137	42	18	60	25.07	6.50
gesamt	223	42	18	60	27.00	8.37

In den beiden folgenden Tabellen 30 und 31 ist die Verteilung der Versuchspersonen nach Beruf, Studium und Geschlecht sowie auf die verschiedenen Fächer dargestellt. Die bei der

Befragung der Seminarteilnehmer der Berufsgenossenschaft aufgetretene Vielzahl der unterschiedlichen Ausbildungsberufe wurde unter der Kategorie Berufstätige zusammengefasst.

Tabelle 31: Aufteilung nach Studium bzw. Beruf und Geschlecht

	Geschlecht	
	männlich	weiblich
Beruf	36	14
Lehramtstudium	21	30
Studium	29	93
Gesamt	86	137

Tabelle 32: Aufteilung nach Fächern und Geschlecht

	Geschlecht	
	Männlich	weiblich
Erziehungswissenschaft	1	5
Geisteswissenschaft	5	1
Heilpädagogik	0	1
Informationswissenschaft	0	1
Mathematik	2	3
Medizin	10	14
Musikwissenschaft	1	0
Naturwissenschaft	2	2
Psychologie	6	62
Rechtswissenschaft	1	1
Sozialwissenschaft	4	2
Sport	0	1
Sportwissenschaft	1	0
Sprachwissenschaft	11	23
Theologie	1	1
Wirtschaftspädagogik	1	0
Wirtschaftswissenschaft	4	1
Berufstätige	35	15
keine Angabe	1	4
Gesamt	86	137

Bei den Angaben zur letzten Mathematiknote sowie der Gesamtnote des Schulabschlusses traten häufig fehlende Werte auf, ebenso wurde die Frage nach der Anzahl der Geschwister häufig nicht beantwortet. Daher werden diese Angaben nicht weiter berücksichtigt. Bei Versuchspersonen, bei denen weniger als sechs fehlende Werte bei den Items des Schommer-Fragebogens vorhanden waren, wurden diese mithilfe des EM-Algorithmus geschätzt (Allison, 2012, 2009, Lüdtke, Robitzsch, Trautwein & Köller, 2007). Die Auswertung erfolgte mithilfe der beiden Programme SPSS 15 sowie Amos 16.

Die Daten wurden auf das Vorliegen von Normalverteilung bzw. multivariater Normalverteilung hin untersucht, da auch eine Auswertung mittels konfirmatorischer Faktorenanalyse erfolgen sollte. Keines der Items wies Schiefen und Wölbungen außerhalb der von West et al. (1995) angegebenen Grenzen auf. Tabelle 32 zeigt die Abweichungen der Items von der Normalverteilung sowie den Mardia-Test auf Abweichung von der multivariaten Normalverteilung. Der Mardia-Test ist signifikant, weshalb standardmäßig die Bollen-Stine-Bootstrap Methode angewandt und die Möglichkeit der Verzerrung der geschätzten Parameter bei der Interpretation der Faktorladungen berücksichtigt wird.

Tabelle 33: Verteilungsform der EBQ-Items

Variable	Schiefe	CR	Wölbung	CR
Item 01	-.82	-4.99	.10	.31
Item 02	-.89	-5.40	.86	2.63
Item 03	-.78	-4.75	.42	1.27
Item 04	-.37	-2.25	-.41	-1.25
Item 05	-.41	-2.48	-.42	-1.26
Item 06	.03	.18	-.70	-2.12
Item 07	-.26	-1.57	-.23	-.71
Item 08	-.56	-3.40	.29	.87
Item 09	-.42	-2.54	-.20	-.60
Item 10	-.65	-3.97	.49	1.49
Item 11	-.21	-1.30	-.55	-1.69
Item 12	.19	1.15	-.58	-1.76
Item 13	.30	1.85	-.37	-1.14
Item 14	-.52	-3.19	.42	1.28
Item 15	-.72	-4.39	.79	2.42
Item 16	-.03	-.18	-.10	-.31
Item 17	-.94	-5.71	1.28	3.90
Item 18	-.74	-4.49	-.18	-.55
Item 19	-.55	-3.34	-.12	-.37
Item 20	.39	2.38	-.40	-1.23
Item 21	-.23	-1.38	-.71	-2.17
Item 22	-.26	-1.59	-.32	-.98
Item 23	-.36	-2.20	-.11	-.32
Item 24	-.51	-3.11	.97	2.95
Item 25	-.70	-4.28	1.36	4.13
Item 26	.10	.64	-.52	-1.58
Item 27	-.33	-1.99	-.46	-1.40
Item 28	-1.00	-6.07	1.33	4.06
Item 29	.32	1.93	-.61	-1.87
Item 30	-.40	-2.44	-.66	-2.01
Item 31	.86	5.21	.36	1.10
Item 32	-1.06	-6.49	1.59	4.84
Item 33	-.37	-2.26	-.49	-1.50
Item 34	.36	2.19	-1.15	-3.50
Item 35	-.89	-5.41	1.27	3.87

Ergebnisse zum EBQ

Variable	Schiefe	CR	Wölbung	CR
Item 36	-.33	-1.99	.20	.62
Item 37	-.41	-2.49	-.34	-1.05
Item 38	-.85	-5.18	1.42	4.32
Item 39	-1.03	-6.27	1.92	5.86
Item 40	-.46	-2.78	.18	.55
Item 41	-.08	-.47	-.08	-.25
Item 42	-.33	-1.99	-1.15	-3.49
Item 43	-1.10	-6.68	1.94	5.91
Item 44	.48	2.91	-.54	-1.63
Item 45	-.60	-3.63	.37	1.12
Item 46	-.34	-2.07	.31	.95
Item 47	-.38	-2.32	-.08	-.24
Item 48	-.87	-5.30	-.39	-1.20
Item 49	.27	1.62	-.58	-1.75
Item 50	1.07	6.52	.86	2.62
Item 51	.06	.38	-.96	-2.92
Item 52	.58	3.52	.25	.77
Item 53	-.93	-5.66	1.16	3.52
Item 54	-.97	-5.93	2.03	6.19
Item 55	.75	4.57	.03	.10
Item 56	.67	4.10	.17	.53
Item 57	-.42	-2.54	-.04	-.12
Item 58	.02	.12	-.65	-1.97
Item 59	-.15	-.92	-.55	-1.68
Item 60	-.80	-4.90	1.44	4.40
Item 61	-.46	-2.82	.75	2.28
Item 62	.53	3.21	-.07	-.21
Item 63	.95	5.80	.54	1.65
Mardia-Test			377.87	31.18

4.1.2 Auswertung nach dem Modell von Schommer

Zu Beginn wird die erste Forschungsfrage geprüft, ob sich die von Schommer (1990) vorgeschlagene Modellvorstellung, die sich in der postulierten Faktorenstruktur des EBQ-Fragebogens wiederfindet, empirisch bestätigen lässt. Dazu wird die Zuordnung der Items zu den von Schommer postulierten Subskalen sowie die Zuordnung dieser Subskalen zu den jeweiligen Faktoren überprüft. Somit müssen insgesamt zwei unterschiedliche Sachverhalte überprüft werden: Zum einen wird mittels konfirmatorischer Faktorenanalyse untersucht, ob sich die von Schommer vorgegebene Item-Skalen-Zuordnung bzw. die Skalen-Faktoren-Zuordnung empirisch bestätigen lässt. Zusätzlich wird mittels einer explorativen Faktorenanalyse der Subskalen der Vergleich mit mehreren Untersuchungen (Schommer, 1990, Schommer et al, 1992, Schommer, 1993a) angestellt, bei denen ebenfalls die Subskalen einer explorativen Faktorenanalyse unterzogen wurden.

Zunächst wird somit die theoretische Zuordnung der Items zu den jeweiligen Subskalen überprüft. Dazu werden jeweils einfaktorielle CFA-Modelle berechnet, um die Modellpassung zu überprüfen. Bei einer mindestens akzeptablen Modellpassung wird zudem noch die Richtung der Faktorladung beurteilt. Diese muss der theoretisch angenommenen Polung des Items entsprechen. In Tabelle 34 sind die Zuordnungen der Items zu den Subskalen zu finden, die Symbole + und – geben die Polung der Items an. Diese Zuordnung ist die Grundlage für die Erstellung der Messmodelle.

Tabelle 34. Zuordnung der Items zu den Subskalen

Skalen Nr.	Name der Skala	Anzahl Items	Items/Polung
S1	Success is unrelated to hard work	4	26-, 32-, 43-, 49+
S2	Avoid ambiguity	5	9+, 27-, 41+, 42+, 44+
S3	Seek single answers	11	2-, 11+, 16+, 17+, 19+, 22-, 23-, 33+, 56-, 58+, 59+
S4	Avoid Integration	8	3+, 14-, 18-, 35+, 37+, 38+, 54-, 63+
S5	Depend on authority	4	1-, 5+, 36+, 40+
S6	Ability to learn is innate	4	8+, 47+, 55+, 57+
S7	Don't criticise authority	6	6+, 7-, 13+, 31+, 45-, 46-
S8	Knowledge is certain	6	12+, 21+, 30-, 34+, 48-, 61-
S9	Learning is quick	5	10+, 29+, 39-, 50+, 60-
S10	Can't learn how to learn	5	4-, 15-, 25-, 28-, 62+
S11	Learn first time	3	20+, 24-, 52+
S12	Concentrated effort is a waste of time	2	51+, 53-

Tabelle 35 enthält die Modellpassungen der Subskalen. Bei den Skalen S11 und S12 ist keine Prüfung möglich, S11 ist ein gerade identifiziertes Messmodell und somit ist keine Modellprüfung möglich. Die Subskala 12 hat nur 2 Items und ist somit nicht identifiziert. Bei den restlichen Skalen zeigten sich nur bei S6 und S10 gute Modellpassungen. Bei den übrigen Skalen erreichten eine oder mehrere Fitindizes nicht die erforderliche Grenze für eine akzeptable Modellpassung.

Tabelle 35: Modellpassung der Subskalen von Schommer

Modell	χ^2 (df)	<i>p</i>	Bollen-Stine- <i>p</i>	χ^2/df	CFI	TLI	RMSE A	CI (90 %)	<i>p_{close}</i>	SRMR
S1*	.99 (2)	.60	<i>n. A.</i>	.50	1.00	1.42	.000	.000;.108	.76	<i>n. A.</i>
S2	14.86 (5)	.01	.03	2.97	.86	.71	.094	.041;.151	.09	.0609
S3	56,28 (44)	.10	.29	1.28	.78	.73	.035	.000;.060	.81	.0570
S4	71,42 (20)	***	***	3.57	.64	.50	.108	.081;.135	***	.0880
S5*	.00 (0)	<i>n. A.</i>	<i>n. A.</i>	<i>n. A.</i>	1.00	<i>n. A.</i>	.066	.000;.120	.27	<i>n. A.</i>
S6	1.48 (2)	.48	.55	.74	1.00	1.02	.000	.000;.122	.64	.0240
S7	46,10 (9)	***	***	5.12	.63	.38	.140	.099;.176	***	.0830
S8	70,01 (9)	***	***	7.78	.43	.05	.181	.138;.214	***	.1245
S9	10.33 (5)	.07	.07	2.07	.81	.62	.069	.000;.0130	.25	.0513
S10	6.17 (5)	.29	.28	1.23	.94	.88	.032	.000;.103	.57	.0393
S11	gerade identifiziert, keine Prüfung der Modellpassung möglich									
S12	unteridentifiziert									

Anmerkung: * Skalen, bei den Schätzprobleme auftraten, ***: $p < .001$

Bei den Skalen S1 und S5 traten Schätzprobleme auf, die auf empirische Unteridentifizierung zurückzuführen sind. Empirische Unteridentifikation bedeutet, dass, obwohl das Modell statistisch gerade oder überidentifiziert ist, die empirische Kovarianzmatrix nicht genügend Information zur Schätzung eines Modells liefert. Dies ist insbesondere dann der Fall, wenn alle oder einige Kovarianzen gleich null sind. Empirische Unteridentifikation führt zu Heywood-Fällen, das sind Parameterschätzungen, die außerhalb des Bereichs plausibler Werte für eine Schätzung liegen, wie bspw. Varianzen mit numerischen Werten von null, negative Varianzen oder standardisierte Regressionsgewichte größer eins (Brown, 2006). Bei der Skala S1 ergibt eine Inspektion der Stichprobenkorrelationsmatrix, dass lediglich zwischen Item 32 und 43 eine Korrelation von .23 vorliegt. Die Stichprobenkorrelationsmatrix für die Skala S1 ist in Tabelle 36 dargestellt.

Tabelle 36: Korrelationsmatrix der Skala S1

	Item 49	Item 43	Item 32	Item 26
Item 49	-			
Item 43	-.05	-		
Item 32	.01	.23	-	
Item 26	.06	-.01	-.03	-

Diese Korrelationen weisen numerische Werte auf, die fast gegen null gehen. Die Schätzung des Modells, bei der die Faktorladung von Item 26 auf eins fixiert wurde, führte zu einer Schätzung der Varianz der latenten Variable von .001 (CR= .12, $p = .90$) und der Fehlervarianz von Item 43 von .077 (CR = .02, $p = .83$), die somit beide nicht signifikant von null ver-

schieden ist. Da Fälle empirischer Unteridentifikation auch auftreten können, wenn die Skaliervariable nicht mit den anderen Variablen korreliert, wurde das Modell ein zweites Mal geschätzt, wobei zur Skalierung die Varianz der latenten Variablen auf eins fixiert wurde. Bei dieser Schätzung wurde die Fehlervarianz von Item 43 negativ geschätzt, was einen Heywood-Fall darstellt. Beide Schätzvarianten führen also zu verzerrten Ergebnissen. Bei diesem Modell kann man somit von einer empirischen Unteridentifikation ausgegangen werden.

Bei Skala S5 zeigt sich, dass die Varianz der latenten Variablen sowie die Fehlervarianz von Item 01 nicht geschätzt werden kann, ebenso die Faktorladungen der Items 05, 36 und 40. Ein Wechsel der Skaliervariablen von Item 01 zur Fixierung der Varianz der latenten Variablen auf eins führte zu einer negativen Schätzung der Fehlervarianz von Item 36, was ebenfalls einen Heywood-Fall darstellt. Eine Inspektion der Stichprobenkorrelationsmatrix ergibt, dass lediglich zwischen den Items 01 und 40, 05 und 36 sowie und 01 und 05 wenigstens ausreichend hohe Korrelationen zu finden sind, alle anderen Korrelationen liegen nahe null. Die Stichprobenkorrelationsmatrix ist in Tabelle 37 dargestellt. Auch bei der Skala S5 kann somit von empirischer Unteridentifikation ausgegangen werden.

Tabelle 37: Korrelationsmatrix der Skala S5

	Item 40	Item 36	Item 05
Item 36	.15	-	
Item 05	.04	.13	-
Item 01	.03	-.03	.11

Beide Skalen lassen sich aufgrund empirischer Unteridentifikation somit nicht schätzen. Die dargestellten Korrelationsmatrizen lassen erkennen, dass nur wenige der den Skalen zugeordneten Items korrelieren, was ein Indiz dafür ist, dass die Items wenig gemeinsame Varianz teilen und somit ein den Items gemeinsamer Faktor nicht gegeben ist.

Wie die Fitindizes aus Tabelle 35 gezeigt haben, findet sich nur bei den Skalen S6 und S10 ein empirischer Beleg für die von Schommer vorgenommene Zuordnung der Items. Die Modelle wiesen einen guten Modellfit auf. Die Schätzungen der Faktorladungen sind in den Tabellen 37 und 38 aufgeführt.

Tabelle 38: Parameterschätzungen S6

Item	partielles Regressionsgewicht		SE	CR	<i>p</i>
	nicht standardisiert	standardisiert			
Item 08	.65	.57	.12	5.35	***
Item 47	1.02	.89	.17	6.13	***
Item 55	.23	.20	.09	2.56	.01
Item 57	.31	.26	.09	3.29	***

Anmerkung: ***: $p < .001$ **Tabelle 39: Parameterschätzungen S10**

Item	partielles Regressionsgewicht		SE	CR	P
	nicht standardisiert	standardisiert			
Item 04	.34	.26	.14	2.44	.02
Item 15	.42	.39	.13	3.18	***
Item 25	.20	.24	.09	2.25	.03
Item 28	.59	.55	.16	3.62	***
Item 62	-.21	-.17	.13	-1.65	.10

Anmerkung: ***: $p < .001$

Bei beiden Messmodellen entsprechen die Richtungen der Faktorladungen der von Schommer (1990) theoretisch postulierten Zuordnung. Allerdings sind nicht alle Faktorladungen signifikant bzw. sind relativ klein. Numerisch kleine Faktorladungen zeigen sich bei den Items 55 und 56 von Skala S6, obwohl beide Faktorladungen signifikant sind. Bei Skala S10 ist die Faktorladung von Item 62 nicht signifikant weist einen sehr kleinen Wert auf. Überdies sind bis auf die Items 15 und 28 bei Skala S10 alle übrigen Faktorladungen auch gering. Die beiden Skalen S6 und S10 entsprechen somit der Zuordnung von Schommer, wenn auch aufgrund der geringen Faktorladungen bei einigen Items die Zuordnung fraglich erscheint.

Obwohl die Messmodell für die zwölf Skalen bis auf zwei Ausnahmen schlechte Modellpassungen aufweisen, wird trotzdem wie vorgesehen das von Schommer (1990) vorgeschlagene Gesamtmodell (vgl Kapitel 1.3.1) mithilfe einer konfirmatorischen Faktorenanalyse untersucht. Dazu wurden die Items gemäß Schommers Angaben umkodiert und anschließend für jede der Subskalen zu einem Summenwert verrechnet. Die Subskalen wurden anschließend den jeweiligen Faktoren zugeordnet (vgl. Schommer, 1990), die Faktorkorrelationen wurden frei geschätzt. Das Modell zeigt folgende Fitindizes (da nicht mit den eigentlichen Daten, sondern Summenwerte gerechnet wurde, konnte kein Bollen-Stine-Bootstrap berechnet werden): $\chi^2 = 84.84$ ($df = 44$), $p < .001$, $\chi^2/df = 1.93$, CFI = .81, TLI = .72, SRMR = .0723, RMSEA = .065 (CI = [.044;.085], $p_{close} = .12$). Das Modell ergibt somit bei einigen Fitindizes akzeptable Werte, aber der CFI und TLI liegen außerhalb des akzeptablen Bereichs, sodass das Modell letztendlich als nicht passend verworfen werden muss.

In Anlehnung an das Vorgehen in den Publikationen von Schommer (1990), Schommer et al. (1992) sowie Schommer (1993a) wird trotz der fehlenden Modellpassung der Schommer-Skalen zusätzlich eine explorative Faktorenanalyse der Subskalen vorgenommen, um diese Ergebnisse mit den Ergebnissen aus den genannten Publikationen vergleichen zu können.

Als Kriterium zur Bestimmung der Faktorzahl wird das Kaiserkriterium (Eigenwert > 1) verwendet, die Lösung wird orthogonal mithilfe einer Varimax-Rotation rotiert. Der Scree-Plot ist in Abbildung 21 dargestellt.

Der Eigenwertverlauf legt eine 3- bis 4-faktorielle Lösung nahe. Aus dem Kaiser-Kriterium erfolgt die Extraktion einer 4-faktoriellen Lösung. Diese klärt 31.2 % der Varianz auf. Das KMO-Maß beträgt .68 und gibt damit an, dass die Daten nicht für die Durchführung einer Faktorenanalyse geeignet sind. Die rotierte Faktorladungsmatrix ist in Tabelle 40 angegeben.

Scree-Plot

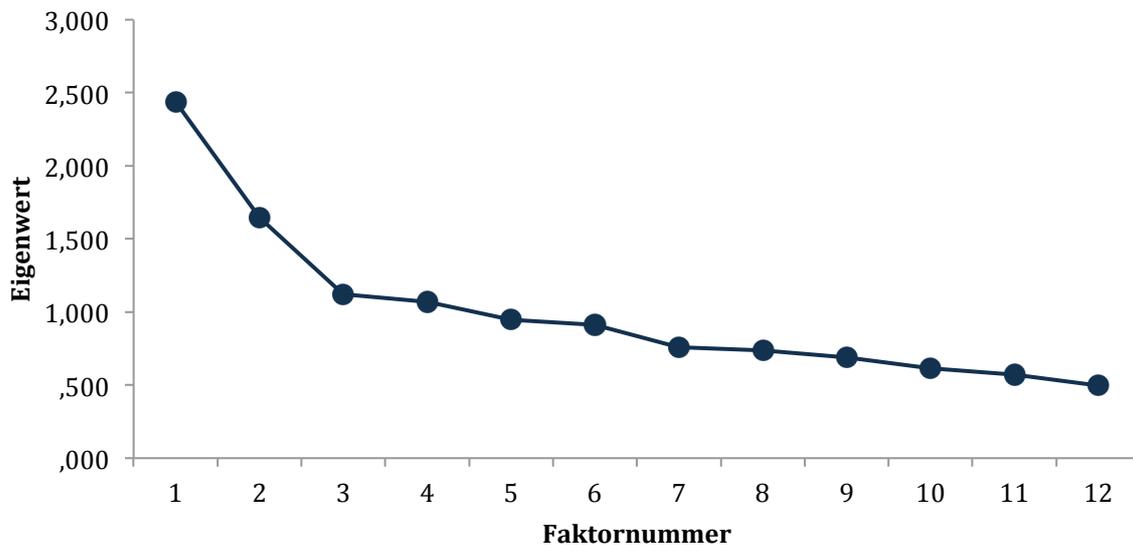


Abbildung 21: Scree-Plot der Schommer Subskalen

Tabelle 40: Faktorladungen der Subskalen

Item	Faktor			
	1	2	3	4
Seek single answers	.53	-.04	-.22	-.14
Don't criticize authority	.51	.28	-.02	.02
Knowledge is certain	.46	-.05	.02	.15
Avoid ambiguity	.45	.39	.02	.05
Ability to learn is innate	.38	.02	-.06	.23
Avoid integration	.31	.30	.07	-.03
Depend on authority	.30	.05	.07	-.01
Learn first time	.03	.51	.09	.06
Concentrated effort is a waste of time	.01	.45	.05	.25
Can't learn how to learn	-.01	.14	.79	.10
Success is unrelated to hard work	.03	.07	.23	.59
Learning is quick	.09	.34	-.13	.49

Nach Schommer (1990) sind nur die Subskalen für die Interpretation eines Faktors relevant, die eine Faktorladung größer als .50 aufweisen. In Tabelle 41 wird die hier gefundene Faktorenstruktur mit den Ergebnissen von Schommer (1990), Schommer et al. (1992) sowie Schommer (1993a) entsprechend der Übersicht nach Clarebout et al. (2001) verglichen, dargestellt werden diejenigen Subskalen, welche eine Faktorladung größer als .40 auf einem der Faktoren besitzen (vgl. Clarebout et al., 2001). Tabelle 40 zeigt, dass das Ladungsmuster nicht ganz mit den 4-faktoriellen Lösungen von Schommer übereinstimmt. Aber alle Subskalen, die eine Ladung über .50 in der hier durchgeführten Studie aufwiesen, wiesen in den vergleichbaren Studien von Schommer ebenso Ladungen über .50 auf.

Die Auswertung des Schommer-Fragebogens auf Grundlage der Subskalen liefert damit zweideutige Ergebnisse. Einerseits weisen die konfirmatorischen Faktorenanalysen der Messmodelle darauf hin, dass diese – bis auf zwei Ausnahmen – nicht der von Schommer postulierten Zuordnung von Items zu den Skalen entsprechen. Auch das Faktormodell auf Basis der Subskalen führte zu keiner ausreichenden Modellpassung. Andererseits lieferte die explorative Faktorenanalyse Ergebnisse, die teilweise in das Muster der Ergebnisse von Schommer passen.

Tabelle 41: Vergleich der Faktorlösung mit den Ergebnissen von Schommer (nach Clarebout et al., 2001)

Untersuchung	Schommer (1990)	Schommer et al. (1992)		Schommer (1993a)	Eigene Untersuchung
		3 Faktoren	4 Faktoren		
N	266	424	424	n. a.	223
Analyse	Hauptfaktoren	Hauptfaktoren	Hauptfaktoren	Hauptfaktoren	Hauptfaktoren
Varianzaufklärung	55.2 %	46.2 %	54.3 %	n. a.	31.1%
Rotation	Varimax	Varimax	Varimax	Varimax	Varimax
Faktorname	F1: Innate Ability F2: Simple knowledge F3: Quick learning F4: Certain knowledge	F1: Innate Ability F2: Simple knowledge F3: Certain knowledge	F1: Externally controlled learning F2: Simple knowledge F3: Quick learning (F4: Certain Knowledge)	F1: Fixed ability F2: Simple Knowledge F3: Quick learning F4: Certain knowledge	
Subskala					
Learning is quick	F3	F1	F3	F3	F4
Can't learn how to learn	F1	F1	F1	F1	F3
Learn first time	F1	F1	F1	F1	F2
Concentrated effort is waste of time		F1	F1		F2
Success is unrelated to hard work	F1	F1	F1	F1	F4
Avoid ambiguity	F2	F2	F2	F2	F1
Seek single answers	F2	F2	F2		F1
Avoid integration		F2	F2	F2	
Depend on authority		F2	F2		
Ability to learn is innate				F3	
Don't criticize authority		F3	(F4, .39)		F1
Knowledge is certain	F4		(F4, .38)	F4	F1

4.1.3 Explorative Faktorenanalyse des Schommer Fragebogens auf Itemebene

Die zweite Forschungsfrage behandelt die Faktorenstruktur des EBQ auf Itemebene. Um diese Forschungsfrage zu untersuchen, werden im Folgenden die erhobenen Daten mithilfe einer explorativen Faktorenanalyse untersucht. Anders als bei der ersten Forschungsfrage steht dabei aber nicht die vorgegebene Zuordnung von Items zu Skalen und von Skalen zu Faktoren im Vordergrund, sondern die Zuordnung von Items zu den Faktoren wird aufgrund der Daten ermittelt. Zuerst wird die Eignung der Daten zur Durchführung einer EFA geprüft und ggf. Maßnahmen zur Verbesserung der Daten durchgeführt, anschließend wird die Anzahl der zu extrahierenden Faktoren bestimmt und letztlich wird die Faktorenlösung zur Interpretation rotiert. Für die explorative Faktorenanalyse werden die gleichen Daten wie im vorhergehenden Kapitel verwendet.

Der erste Schritt besteht in der Prüfung der Eignung der Daten zur Durchführung einer EFA. Das KMO-Maß für alle 63 Items betrug $.64$ und ist damit nur als mäßig zu bezeichnen. Um ein differenzierteres Bild über die Eignung der Daten zu erhalten, wurde eine Inspektion der einzelnen MSA-Koeffizienten durchgeführt. Es zeigten sich MSA-Werte, die teilweise unter $.60$ waren und damit als schlecht zu bezeichnen sind, MSA-Werte unterhalb von $.50$ gelten als nicht vereinbar mit der Durchführung einer Faktorenanalyse. Im ersten Schritt wurden daher diejenigen Items ausgewählt, deren MSA größer als $.60$ war. Dadurch sollte sichergestellt werden, dass nur solche Items in die Faktorenanalyse aufgenommen werden, die minimal miteinander zusammenhängen. Zwar ist das Kriterium $MSA > .60$ recht liberal, ein strengeres Kriterium würde ggf. zu einer zu starken Einschränkung der Items führen. Nach der ersten Elimination von Items mit einem $MSA < .60$ wurden die MSA-Werte der einzelnen Items erneut berechnet. Dabei wies das Item 58 einen MSA-Wert kleiner als $.60$ auf und wurde ausgeschlossen. Die anschließende erneute Berechnung der MSA-Werte ergab, dass jedes Item einem $MSA > .60$ aufwies. Es wurden keine weiteren Items ausgeschlossen. Insgesamt verblieben nach dem Ausschluss 38 Items. Die Items sowie die MSA-Werte sind in Tabelle 42 aufgeführt. Durch den Ausschluss der Items verbesserte sich das KMO-Maß auf $.73$. Nach den Bewertungskriterien für das KMO-Maß sind sie damit im Bereich einer mittelmäßigen Eignung für eine Faktorenanalyse.

Tabelle 42: MSA-Werte der Schommer-Items

Item	Formulierung	MSA 1	MSA 2	MSA 3
1	Wenn in einem Lehrtext ein schwer verständlicher Begriff gebraucht wird, ist es am besten, diesen zunächst einmal zu klären, bevor man weiter liest.	.53		
2	Ein Satz hat wenig Bedeutung, solange man den Kontext, in dem er gesprochen oder geschrieben wurde, nicht kennt.	.58		
3	Ein guter Student zu sein, bedeutet, auch Fakten auswendig zu lernen.	.58		
4	Ein Kurs in Lern- und Studiertechniken wäre hilfreich für mich.	.56		
5	Wie viel man lernt, hängt in erster Linie von den Dozenten ab.	.50		
6	Man kann fast alles glauben, was man für das Studium liest.	.77	.76	.75
7	Ich wundere mich oft, wie viel meine Dozenten wissen.	.64	.67	.66
8	Unterschiede in der Lernfähigkeit sind angeboren.	.47		
9	Es ärgert mich, wenn der Dozent nicht klar sagt, was er zu einem Problem denkt.	.67	.68	.68
10	Erfolgreiche Studenten verstehen Lerninhalte schnell.	.65	.66	.66
11	Ein guter Dozent muss dafür sorgen, dass die Studenten nicht von der "richtigen Spur" abkommen.	.63	.70	.70
12	Wenn Wissenschaftler hart arbeiten, können sie über fast alles die Wahrheit herausfinden.	.68	.70	.70
13	Leute, die Autoritäten anzweifeln, überschätzen sich.	.74	.81	.80
14	Ich versuche, möglichst weitgehend die Lerninhalte verschiedener Stunden und verschiedener Lernveranstaltungen zu integrieren.	.72	.74	.73
15	Die erfolgreichsten Leute haben entdeckt, wie sie ihre Lernfähigkeit verbessern können.	.68	.72	.72
16	Die Dozenten machen die Dinge komplizierter als sie sind.	.50		
17	Der wichtigste Aspekt wissenschaftlichen Arbeitens ist präzise Messung und Sorgfalt.	.62	.78	.78
18	Einen Text durcharbeiten heißt für mich weniger einzelne Fakten zu lernen, als die Hauptideen zu erfassen.	.74	.74	.75
19	Die Dozenten sollten wissen, was die beste Lehrmethode ist.	.59		
20	Ein schwieriges Kapitel immer und immer wieder zu lesen, hilft wenig, es zu verstehen.	.50		
21	Wissenschaftler können letztendlich zur Wahrheit kommen.	.67	.67	.69

Ergebnisse zum EBQ

Item	Formulierung	MSA 1	MSA 2	MSA 3
22	Man weiß nie, was ein Buch besagt, solange man die Absichten eines Autors nicht kennt.	.55		
23	Der wichtigste Aspekt wissenschaftlichen Arbeitens sind originelle Ideen.	.53		
24	Wenn ich die Zeit finde, einen Lehrtext noch einmal zu lesen, lerne ich beim zweiten Mal viel mehr.	.72	.71	.71
25	Es liegt am jeweiligen Studenten, wie viel er aus einem Lehrbuch lernt.	.63	.67	.68
26	Genialität hat mehr mit harter Arbeit als mit Intelligenz zu tun.	.60		
27	Es macht mir Spaß, über Dinge nachzudenken, über die sich Fachleute nicht einigen können.	.69	.70	.70
28	Jeder muss lernen, wie man lernt.	.60		
29	Wenn man etwas überhaupt verstehen kann, erscheint dies einem schon beim ersten Mal als sinnvoll.	.74	.80	.79
30	Das einzig Gewisse ist die Ungewissheit.	.66	.66	.66
31	Wenn man an der Universität erfolgreich sein will, stellt man am besten nicht allzu viele Fragen	.69	.69	.72
32	Weisheit heißt nicht, die Antworten zu kennen, sondern zu wissen, wie man sie findet.	.69	.72	.73
33	Die meisten Wörter haben eine klare Bedeutung.	.63	.70	.69
34	Wahrheit ändert sich nicht.	.68	.72	.72
35	Wenn man nach der Lektüre eines Textes zwar Einzelheiten vergessen hat, aber zu neuen Ideen gekommen ist, so ist dies ein Lernerfolg.	.71	.79	.79
36	Bei schwierigen Entscheidungen würde ich es am liebsten haben, wenn jemand mir sagen könnte, was richtig ist.	.48		
37	Definitionen Wort für Wort auswendig zu lernen, ist oft nötig, um eine Prüfung gut zu bestehen.	.55		
38	Wenn ich lerne, präge ich mir jeweils spezifische Fakten ein.	.58		
39	Wenn man etwas nicht gleich versteht, sollte man erst recht dranbleiben.	.77	.77	.77
40	Manchmal muss man die Antworten des Dozenten akzeptieren, auch wenn man sie nicht versteht.	.56		
41	Wenn die Dozenten sich mehr auf Fakten als auf Theoretisieren stützen würden, würde man mehr von ihren Veranstaltungen profitieren.	.69	.70	.71
42	Ich mag Filme mit offenem Ende nicht.	.62	.64	.63

Ergebnisse zum EBQ

Item	Formulierung	MSA 1	MSA 2	MSA 3
43	Vorankommen im Studium heißt viel arbeiten.	.74	.75	.74
44	Es ist Zeitverschwendung, sich mit Problemen zu beschäftigen, bei denen man nicht zu einer klaren und eindeutigen Antwort kommen kann.	.73	.75	.77
45	Man sollte Lehrtexte immer hinterfragen.	.60		
46	Auch der Rat von Fachleuten ist oft anzuzweifeln.	.61	.73	.72
47	Einige werden als gute Lerner geboren, andere haben diesbezüglich nur begrenzte Möglichkeiten.	.66	.64	.66
48	Nichts außer dem Tod ist sicher.	.63	.69	.67
49	Wirklich gute Studenten müssen nicht viel arbeiten, um erfolgreich ihr Studium zu bestreiten.	.59		
50	Es zahlt sich nur für die klügsten Studenten aus, sich intensiv längere Zeit mit einem Problem zu beschäftigen.	.78	.80	.80
51	Wenn man zu intensiv versucht, ein Problem zu verstehen, ist man am Ende nur verwirrt.	.64	.71	.72
52	Man lernt fast alles, was man aus einem Lehrtext lernen kann, beim ersten Lesen.	.66	.68	.69
53	Ich verstehe schwierige Sachverhalte dann, wenn ich alle Störungen von außen abschirme und mich wirklich konzentriere.	.54		
54	Eine gute Art, einen Lehrtext zu verstehen, ist es, die Informationen entsprechend dem eigenen Vorwissen neu zu ordnen.	.68	.76	.77
55	Schüler, die in der Schule durchschnittlich sind, werden im ganzen Leben "Durchschnitt" bleiben.	.77	.82	.82
56	Ein geordneter Kopf ist ein leerer Kopf.	.58		
57	Um ein Experte auf einem Gebiet zu werden, muss man ein besonderes Talent dafür mitbringen.	.57		
58	Ich schätze es, wenn Dozenten ihre Lehrveranstaltungen minutiös planen und sich dann an ihren Plan halten.	.62	.60	
59	Das Beste an der Wissenschaft ist, das sie zeigt, dass es auf die meisten Probleme eine richtige Antwort gibt.	.69	.75	.75
60	Lernen ist ein langsamer Prozess des Wissensaufbaus.	.55		
61	Derzeit gültige Fakten sind meist schon bald veraltet.	.61	.64	.67
62	Bücher zur Selbsthilfe sind von geringem Nutzen.	.60		

Item	Formulierung	MSA 1	MSA 2	MSA 3
63	Es verwirrt nur, wenn man die Inhalte eines Lehrtextes mit dem in Verbindung bringt, was man bereits weiß.	.69	.73	.72

Anmerkung: MSA1: MSA-Werte bei der ersten Berechnung, MSA2: MSA-Werte nach dem ersten Ausschluss, MSA3: MSA-Werte nach Ausschluss von Item 58.

Als Nächstes muss die Anzahl der zu extrahierenden Faktoren bestimmt werden. Dabei wird die Parallelanalyse nach Horn (1965) sowie das von Costello und Osborne (2005) vorgeschlagene Verfahren zur Bestimmung der Faktorzahl verwendet.

Die Ergebnisse der Parallelanalyse für die 38 verbleibenden Variablen sind in Tabelle 43 aufgeführt. Es wurden 100 Zufallsdatensätze durch Permutation der ursprünglichen Daten der 223 Versuchspersonen erzeugt und für diese die Eigenwertverläufe berechnet. In der Tabelle sind die Mittelwerte sowie die 95%-Perzentile für die Verteilungen der Eigenwerte der 100 Zufallsdatensätze enthalten. Bis zum sechsten Eigenwert sind die Eigenwerte der Stichprobe größer als die Eigenwerte der Zufallsdaten. Die Parallelanalyse legt somit die Extraktion einer 6-faktoriellen Lösung nahe.

Tabelle 43: Ergebnisse der Parallelanalyse

Faktor	Eigenwert Mittel	Eigenwert 95 % Perzentil	Eigenwert Daten
1	1.06	1.18	3.92
2	.96	1.03	3.32
3	.87	.95	1.82
4	.80	.87	1.06
5	.74	.79	1.02
6	.68	.73	.82

Der Scree-Plot weist dagegen auf eine 3-faktorielle Lösung hin, wenn man der Richtlinie von Costello und Osborne (2005) folgt, bzw. auf eine 4- oder 5-faktorielle Lösung, wenn man die Richtlinien aus Bortz (1999) beachtet.

Um den Bereich der möglichen Lösungen hinsichtlich ihrer inhaltlichen Adäquatheit zu evaluieren (vgl. Costello & Osborne, 2005), wurden die 3- bis 6-faktoriellen Lösungen extrahiert. Als Extraktionsmethode wurde die Hauptachsenanalyse mit anschließender Promax-Rotation ($\kappa = 4$) benutzt. Zur inhaltlichen Evaluation wurde die Mustermatrix benutzt.

Scree-Plot

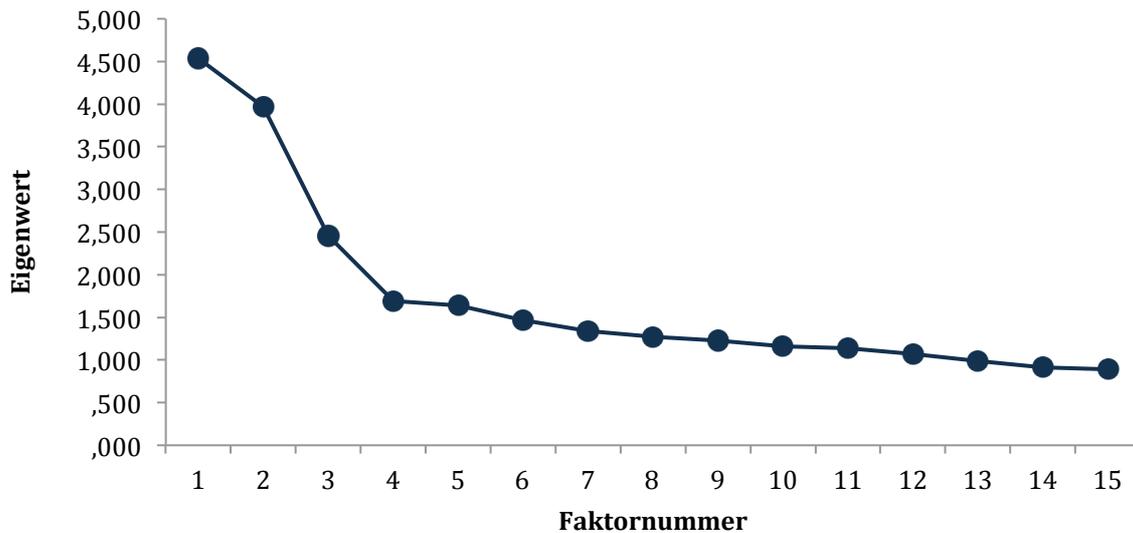


Abbildung 22: Scree-Plot der Schommer Items

Bei den 3- und 4-faktoriellen Lösungen ergaben sich Faktoren, die sich inhaltlich auf verschiedene Themenbereiche bezogen. Bei der 6-faktoriellen Lösung erwies sich der sechste Faktor aus nur einem Item bestehend, sodass nicht von einem stabilen Faktor ausgegangen werden kann. Lediglich die 5-faktorielle Lösung wies inhaltlich eine thematisch getrennte und eindeutig gegliederte Struktur auf.

An die Faktorenlösung wurde die Anforderung gestellt, dass sich jedes Item eindeutig und bedeutsam auf einem Faktor laden sollte. Dazu soll jedes Item eine Ladung $> .30$ auf dem jeweiligen Faktor haben (vgl. Bortz, 1999). In einem weiteren Analyseschritt wurden daher Items, die diese Forderung nicht erfüllen, ausgeschlossen. Weiterhin sollte eine Einfachstruktur vorliegen, d. h. jedes Item sollte nur auf dem jeweils zugeordneten Faktor eine Ladung $> .30$ aufweisen.

Die initiale 5-faktorielle Lösung wies auf dem ersten und dritten Faktor Items auf, die keine Ladung über $.30$ hatten und somit nicht eindeutig dem Faktor zuzuschreiben waren. Auf dem ersten Faktor waren dies die Items 7, 11 und 47 und auf dem dritten Faktor das Item 17. Diese Items wurden für eine weitere Faktorenanalyse ausgeschlossen und es wurde eine neue Hauptachsenanalyse mit anschließender Promax-Rotation ($\kappa = 4$) berechnet. Das KMO-Maß für diese weitere Analyse betrug $.73$. Bei dieser Lösung hatte das Item 41 keine Ladung über $.30$ und wurde deswegen ausgeschlossen. Es wurde nun erneut eine Hauptachsenanalyse ge-

rechnet, das KMO-Maß betrug .73. In dieser Lösung war das Item 29 keinem Faktor zuzuordnen, es wurde also entfernt und eine neue Hauptachsenanalyse wurde berechnet, es ergab sich ein KMO-Maß von .72. In dieser Lösung war nun jedes Item eindeutig einem Faktor zuzuordnen.

Eine Inspektion dieser Lösung ergab, dass jeweils die Items 42 und 48, die ihre Hauptladung auf dem 5. Faktor hatten, auch eine Nebenladung $> .30$ auf dem ersten Faktor aufwiesen. Da dies die beiden einzigen Items mit einer Nebenladung waren, wurden keine weiteren Modifikationen an der Faktorenlösung vorgenommen. Diese finale Lösung ist in Tabelle 44 dargestellt.

Tabelle 44: Mustermatrix der Promax-Rotation

Item Nr.	Faktor					Kommunalität
	F1	F2	F3	F4	F5	
31	.60	-.06	-.10	.03	.16	.35
63	.56	-.01	-.10	-.02	-.01	.34
52	.49	.00	.00	-.13	.27	.28
13	.47	.22	.05	.10	-.01	.33
50	.45	.04	.00	-.24	.00	.31
51	.44	-.02	.29	-.16	-.18	.32
44	.44	.20	.19	-.14	-.23	.41
21	-.06	.62	.10	-.02	.03	.37
12	.15	.49	.18	-.10	.20	.33
33	-.12	.46	.02	.07	-.06	.22
6	.20	.44	-.08	.09	-.11	.34
59	.25	.40	-.28	.20	.05	.38
34	.14	.35	-.06	.02	-.11	.21
55	.28	.35	-.12	-.07	.02	.28
61	.12	-.17	.73	.01	-.07	.53
35	-.11	-.06	.45	.15	.10	.33
10	-.10	.28	.40	-.01	.13	.28
18	-.03	-.03	.38	.01	.23	.26
9	.17	.10	.37	.09	-.04	.18
54	-.13	.05	.36	.24	.02	.26
24	-.13	-.07	.10	.63	-.11	.45
43	-.06	.16	.07	.49	-.03	.30
25	-.02	.01	-.04	.47	.07	.24
39	-.03	.13	.12	.39	.07	.24
32	-.02	-.03	.27	.37	.08	.29
27	.03	-.11	-.02	.07	.57	.36
15	.03	.24	.14	-.01	.44	.27
14	-.29	.21	.15	.01	.42	.37
46	.14	-.28	.16	-.16	.40	.31
42	.34	.02	.14	.23	-.36	.28
48	.32	-.10	-.07	.27	.34	.25
30	.28	-.27	.14	.18	.31	.28

Die Korrelationsmatrix der Faktoren ist in Tabelle 45 wiedergegeben. Die höchste Korrelation ergibt sich zwischen Faktor 3 und 5, gefolgt von der Korrelation der Faktoren 1 und 2.

Tabelle 45: Korrelationsmatrix der fünf Faktoren

Faktor	1	2	3	4
2	.27**	-		
3	-.10	-.03	-	
4	-.14*	.11	.21**	-
5	-.20**	-.16*	.32**	.23**

Anmerkung: *: $p < .05$, **: $p < .001$

Dieses mittels der EFA gewonnene Faktormodell wird mittels einer E/CFA näher auf seine Modellpassung hin untersucht. Dazu wird die vollständige Faktorenlösung in einem CFA-Modell umgesetzt, als Skalierer der latenten Variablen wird jeweils das erste Item eines Faktors gewählt. Der Skalierer wird jeweils nur einem Faktor zugeordnet, während für alle anderen Variablen alle möglichen Kreuzladungen modelliert werden, auch werden alle Kovarianzen (bzw. Korrelationen) der Faktoren modelliert (vgl. Brown, 2006).

Eine Inspektion der Verteilungsform der Items ergab, dass keine der in der EFA verbleibenden Variablen eine Schiefe und Wölbung außerhalb der Grenzwerte von West et al. (1995) aufwies. Der Mardia-Test hingegen wies auf eine Abweichung der Daten von der multivariaten Normalverteilung hin ($M = 173.17$, $CR = 27.72$).

Bis auf Ausnahme des p-Wertes sowie des CFI und TLI deuten die Fitindizes des E/CFA-Modells auf eine akzeptable bis gute Modellpassung hin ($\chi^2 = 551.98$ ($df = 346$), $p < .001$, Bollen-Stine- $p = .02$, $\chi^2/df = 1.59$, $CFI = .84$, $TLI = .77$, $SRMR = .048$, $RMSEA = .052$, $CI = [.044; .060]$, $p_{close} = .35$). Grundsätzlich scheint sich somit die durch EFA ermittelte Lösung auch unter konfirmatorischen Gesichtspunkten zu bewähren. Nur der CFI und TLI lassen auf eine unzureichende Modellpassung schließen. Dies kann aber darin begründet liegen, dass der CFI und TLI sparsamere Modelle bevorzugt und weniger sparsame abstrafte (Schermele-Engel et al., 2003), ähnlich dem NNFI von Bentler und Bonnet (1980), auf welchem der CFI und der TLI basieren. Bei der Modellierung eines EFA-Modells mithilfe eines CFA-Modells werden aber fast alle möglichen Ladungen der manifesten auf den latenten Variablen berücksichtigt und das Modell ist daher nicht als sparsam zu bezeichnen. Daher ist es denkbar, dass Fitindizes wie der CFI die eigentliche Modellpassung nicht richtig widerspiegeln.

Ergebnisse zum EBQ

Tabelle 46: Faktorladungen des E/CFA-Modells

Item	→	Faktor	partielles Regressionsge- wicht		SE	CR	p
			nicht stan- dardisiert	standardi- siert			
Faktorladungen F1							
Item 31	→	F1	1.00	.60			
Item 63	→	F1	.87	.58	.15	5.65	***
Item 52	→	F1	.61	.41	.14	4.47	***
Item 13	→	F1	.86	.52	.17	5.10	***
Item 50	→	F1	.68	.42	.15	4.56	***
Item 51	→	F1	.66	.38	.16	4.13	***
Item 44	→	F1	.74	.42	.17	4.43	***
Faktorladungen F2							
Item 21	→	F2	1.00	.60			
Item 12	→	F2	.83	.52	.16	5.18	***
Item 33	→	F2	.70	.44	.15	4.66	***
Item 06	→	F2	.67	.42	.15	4.46	***
Item 59	→	F2	.55	.36	.14	3.90	***
Item 34	→	F2	.74	.33	.20	3.72	***
Item 55	→	F2	.49	.32	.14	3.55	***
Faktorladungen F3							
Item 61	→	F3	1.00	.80			
Item 35	→	F3	.42	.36	.10	4.21	***
Item 10	→	F3	.40	.31	.12	3.34	***
Item 18	→	F3	.44	.30	.13	3.49	***
Item 09	→	F3	.42	.31	.12	3.60	***
Item 54	→	F3	.29	.26	.09	3.20	***
Faktorladungen F4							
Item 24	→	F4	1.00	.72			
Item 43	→	F4	.63	.46	.14	4.45	***
Item 25	→	F4	.62	.48	.13	4.63	***
Item 39	→	F4	.46	.35	.13	3.51	***
Item 32	→	F4	.51	.33	.15	3.51	***
Faktorladungen F5							
Item 27	→	F5	1.00	.64			
Item 15	→	F5	.60	.46	.14	4.33	***
Item 14	→	F5	.66	.58	.13	4.99	***
Item 46	→	F5	.46	.41	.11	4.24	***
Item 42	→	F5	-.72	-.35	.21	-3.42	***
Item 48	→	F5	.49	.26	.18	2.70	.01
Item 30	→	F5	.42	.25	.16	2.71	.01
Einzelladungen der Items auf den Faktoren							
Item 63	→	F2	.00	.00	.14	.02	.98
Item 63	→	F3	-.05	-.03	.13	-.37	.72
Item 63	→	F4	.00	.00	.18	-.02	.98
Item 63	→	F5	-.23	-.17	.14	-1.67	.10

Ergebnisse zum EBQ

Item	→	Faktor	partiell es Regressionsge- wicht		SE	CR	p
			nicht stan- dardisiert	standardi- siert			
Item 52	→	F5	.17	.13	.13	1.37	.17
Item 52	→	F4	-.38	-.23	.16	-2.32	.02
Item 52	→	F3	.22	.16	.12	1.86	.06
Item 52	→	F2	.11	.08	.13	.85	.39
Item 13	→	F2	.46	.28	.16	2.94	***
Item 13	→	F3	.06	.04	.14	.45	.65
Item 13	→	F4	.21	.11	.20	1.05	.29
Item 13	→	F5	-.14	-.10	.15	-.94	.35
Item 50	→	F5	-.20	-.14	.14	-1.47	.14
Item 50	→	F4	-.41	-.22	.18	-2.27	.02
Item 50	→	F3	.08	.05	.13	.64	.52
Item 50	→	F2	.06	.04	.14	.42	.68
Item 51	→	F2	.04	.02	.15	.25	.81
Item 51	→	F3	.58	.36	.15	3.86	***
Item 51	→	F4	-.27	-.14	.19	-1.42	.16
Item 51	→	F5	-.45	-.29	.16	-2.92	***
Item 44	→	F5	-.52	-.33	.16	-3.30	***
Item 44	→	F4	-.23	-.12	.20	-1.15	.25
Item 44	→	F3	.29	.18	.15	1.99	.05
Item 44	→	F2	.36	.21	.16	2.31	.02
Item 12	→	F1	.27	.16	.16	1.69	.09
Item 33	→	F1	-.10	-.06	.15	-.63	.53
Item 06	→	F1	.41	.25	.15	2.72	.01
Item 59	→	F1	.54	.35	.15	3.68	***
Item 34	→	F1	.44	.19	.20	2.17	.03
Item 55	→	F1	.48	.31	.14	3.39	***
Item 12	→	F3	.23	.15	.14	1.67	.09
Item 33	→	F3	-.11	-.08	.13	-.86	.39
Item 06	→	F3	-.15	-.10	.13	-1.09	.28
Item 59	→	F3	-.38	-.27	.13	-2.91	***
Item 34	→	F3	-.31	-.15	.18	-1.70	.09
Item 55	→	F3	-.14	-.10	.12	-1.16	.25
Item 55	→	F4	-.14	-.08	.17	-.80	.42
Item 34	→	F4	.06	.02	.25	.23	.82
Item 59	→	F4	.31	.18	.18	1.74	.08
Item 06	→	F4	.13	.07	.19	.70	.48
Item 33	→	F4	.12	.07	.18	.66	.51
Item 12	→	F4	-.13	-.07	.19	-.66	.51
Item 12	→	F5	.16	.11	.15	1.09	.28
Item 33	→	F5	-.10	-.07	.14	-.69	.49
Item 06	→	F5	-.32	-.21	.15	-2.18	.03
Item 59	→	F5	-.15	-.10	.14	-1.06	.29
Item 34	→	F5	-.30	-.14	.19	-1.55	.12

Ergebnisse zum EBQ

Item	→	Faktor	partiell Regressionen- gewicht		SE	CR	p
			nicht stan- dardisiert	standardi- siert			
Item 55	→	F5	-.19	-.13	.13	-1.40	.16
Item 35	→	F1	-.18	-.14	.10	-1.70	.09
Item 10	→	F1	-.22	-.16	.13	-1.67	.10
Item 18	→	F1	-.12	-.08	.14	-.91	.36
Item 09	→	F1	.23	.16	.13	1.84	.07
Item 54	→	F1	-.13	-.11	.10	-1.34	.18
Item 54	→	F2	.17	.15	.10	1.82	.07
Item 09	→	F2	.30	.21	.12	2.46	.01
Item 18	→	F2	.13	.09	.13	1.01	.31
Item 10	→	F2	.53	.39	.13	4.12	***
Item 35	→	F2	.09	.07	.10	.86	.39
Item 35	→	F4	.22	.15	.13	1.72	.09
Item 10	→	F4	-.10	-.06	.16	-.61	.55
Item 18	→	F4	.02	.01	.17	.14	.89
Item 09	→	F4	.16	.10	.16	1.05	.29
Item 54	→	F4	.35	.26	.12	2.83	.01
Item 54	→	F5	.16	.14	.10	1.65	.10
Item 09	→	F5	.03	.03	.12	.28	.78
Item 18	→	F5	.49	.34	.13	3.67	***
Item 10	→	F5	.30	.23	.13	2.37	.02
Item 35	→	F5	.29	.25	.10	2.89	***
Item 43	→	F1	.02	.02	.11	.22	.83
Item 25	→	F1	.08	.07	.10	.76	.45
Item 39	→	F1	.02	.02	.10	.24	.81
Item 32	→	F1	-.01	.00	.11	-.05	.96
Item 32	→	F2	.15	.11	.11	1.33	.18
Item 39	→	F2	.28	.24	.10	2.76	.01
Item 25	→	F2	.09	.09	.09	1.00	.32
Item 43	→	F2	.28	.24	.10	2.79	.01
Item 43	→	F3	.02	.02	.09	.22	.83
Item 25	→	F3	-.08	-.08	.09	-.93	.35
Item 39	→	F3	.03	.03	.09	.31	.76
Item 32	→	F3	.29	.23	.10	2.81	.01
Item 32	→	F5	.24	.19	.11	2.21	.03
Item 39	→	F5	.20	.18	.10	1.99	.05
Item 25	→	F5	.10	.10	.09	1.10	.27
Item 43	→	F5	.04	.04	.10	.43	.67
Item 15	→	F4	-.11	-.06	.17	-.62	.54
Item 15	→	F3	.16	.12	.12	1.29	.20
Item 15	→	F2	.47	.33	.14	3.47	***
Item 15	→	F1	.00	.00	.14	.03	.98
Item 14	→	F1	-.40	-.32	.13	-3.04	***
Item 14	→	F2	.36	.30	.13	2.88	***

Ergebnisse zum EBQ

Item	→	Faktor	partiell Regressionen- gewicht		SE	CR	p
			nicht stan- dardisiert	standardi- siert			
Item 14	→	F3	.07	.06	.11	.60	.55
Item 14	→	F4	-.06	-.05	.16	-.40	.69
Item 46	→	F4	-.33	-.24	.13	-2.48	.01
Item 46	→	F3	.32	.29	.10	3.25	***
Item 46	→	F2	-.20	-.17	.10	-1.97	.05
Item 46	→	F1	.02	.01	.11	.16	.87
Item 42	→	F1	.81	.36	.22	3.71	***
Item 42	→	F2	.17	.08	.20	.84	.40
Item 42	→	F3	.20	.10	.19	1.05	.30
Item 42	→	F4	.56	.22	.27	2.11	.04
Item 48	→	F4	.52	.22	.24	2.22	.03
Item 48	→	F3	.06	.03	.17	.37	.71
Item 48	→	F2	.01	.00	.18	.05	.96
Item 48	→	F1	.66	.31	.20	3.40	***
Item 30	→	F1	.43	.23	.17	2.60	.01
Item 30	→	F2	-.25	-.14	.16	-1.61	.11
Item 30	→	F3	.52	.30	.15	3.46	***
Item 30	→	F4	.31	.15	.20	1.58	.12

Anmerkung: ***: $p < .001$

In Tabelle 47 sind die durch die E/CFA geschätzten Korrelationen der Faktoren aufgeführt. In Tabelle 48 sind die Korrelationen der Faktoren der EFA und der E/CFA zum Vergleich zusammen aufgeführt, unterhalb der Hauptdiagonalen finden sich die Faktorkorrelationen der EFA, oberhalb die der E/CFA.

Tabelle 47: Korrelation der Faktoren in der E/CFA

Faktor	1	2	3	4
2	.02			
3	-.11	.00		
4	-.28**	.04	.22**	
5	.10	-.22**	.12	.22**

Anmerkung: **: $p < .001$

Tabelle 48: Vergleich der Faktorkorrelation von EFA und E/CFA

Faktor	1	2	3	4	5
1		.02	-.11	-.28	.10
2	.27		.00	.04	-.22
3	-.10	-.03		.22	.12
4	-.14	.11	.21		.22
5	-.20	-.16	.32	.23	

Anmerkung: unterhalb der Hauptdiagonalen finden sich die Faktorkorrelationen der EFA, oberhalb die der E/CFA

4.1.4 Psychometrische Eigenschaften und inhaltliche Interpretation der Faktoren

Die nächste Forschungsfrage, bezieht sich auf die psychometrischen Eigenschaften der Skalen, die aus den im vorherigen Abschnitt extrahierten Faktoren gebildet werden können. Bei der Konstruktion der Skalen wird zudem ein Schritt nachgeholt, der in der faktorenanalytischen Auswertung bis jetzt nicht berücksichtigt wurde: die inhaltliche Interpretation der aus den Faktoren entwickelten Skalen.

Zur Entwicklung der Skalen gehört neben der eigentlichen Bestimmung der Reliabilität auch eine Überprüfung der Annahmen der Eindimensionalität und der τ -Äquivalenz der Messungen. Nach der in Kapitel 3.3 beschriebenen Vorgehensweise werden dazu einfaktorielle CFA-Modelle untersucht. Dies liefert nicht nur die notwendigen Hinweise auf die Eindimensionalität und Messäquivalenz, sondern es können ggf. auch nicht passende Items aus den Modellen entfernt werden. Hinweise auf nicht passende Items liefern die Fitindizes zusammen mit den Modifikationsindices und der Matrix der standardisierten Residuen. Weist ein eindimensionales Messmodell eine schlechte Modellpassung auf, so kann dies einerseits daran liegen, dass ein oder mehrere Items nicht durch einen gemeinsamen Faktor bedingt werden. Andererseits können mehrere Items korrelierende Fehlervariablen besitzen, was auf einen Methodenfaktor hindeuten kann. Die Überprüfung der Äquivalenzannahmen erfolgt mittels der in Kapitel 3.4.2 dargestellten Restriktionen, die im Messmodell angewandt wurden. Im Anschluss werden die Skalen inhaltlich interpretiert und einer Bezeichnung versehen.³⁷

4.1.4.1 Skala F1: Autoritäres-einfaches Wissen

Der Faktor F1 besteht insgesamt aus sieben Items (Nr. 31, 63, 52, 13, 50, 51, 44). Keine der Variablen zeigte eine merkliche Abweichung von der Normalverteilung, alle Schiefen und Wölbungen waren in den Grenzen von West et al. (1995). Das Modell wies eine gute Modellpassung auf ($\chi^2 = 8.35$ ($df = 14$), $p = .87$, Bollen-Stine- $p = .97$, $\chi^2/df = .59$, CFI = 1.00, TLI = 1.04, SRMR = .0268, RMSEA = .000, CI = [.000;.034], $p_{close} = .99$). Die Faktorladungen der jeweiligen Items sind in Tabelle 49 angegeben. Alle Faktorladungen sind signifikant³⁸ und substantiell.

³⁷ Zur besseren Übersicht wird der Faktornamen in den Unterüberschriften vorher aufgeführt.

³⁸ Die Faktorladung des Skalierers kann nicht auf Signifikanz geprüft werden, da diese auf eins fixiert wurde. Zur Berechnung des standardisierten Regressionsgewichts (Faktorladung) wird das Modell unter Fixierung der Varianz der latenten Variablen auf eins erneut berechnet und die Standardisierung vorgenommen.

Tabelle 49: Faktorladungen des Faktors Autoritäres-einfaches Wissen)

Item	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
Item 31	1.00	.55			
Item 63	1.00	.60	.17	5.84	***
Item 52	.67	.42	.15	4.61	***
Item 13	.94	.52	.18	5.35	***
Item 50	.91	.51	.17	5.29	***
Item 51	.89	.47	.18	5.01	***
Item 44	1.05	.55	.19	5.56	***

Anmerkung: ***: $p < .001$

Die gebildete Skala kann somit als eindimensional gelten. Im nächsten Schritt muss die Äquivalenz der Messung bestimmt werden. Dazu werden entsprechende einfaktorielle CFA-Modelle mit den jeweiligen Restriktionen berechnet. Das kongenerische Modell entspricht dabei dem schon berechneten CFA-Modell ohne die für die τ -Äquivalenz bzw. Parallelität notwendigen Restriktionen.

Für das τ -äquivalente Messmodell ergaben sich folgende Fitindizes: $\chi^2 = 14.70$ ($df = 20$), $p = .79$, Bollen-Stine- $p = .94$, $\chi^2/df = .74$, CFI = 1.00, TLI = 1.03, SRMR = .0423, RMSEA = .000, CI = [.000;0.040], $p_{close} = .98$. Das Messmodell kann somit als τ -äquivalent gelten.

Für das parallele Messmodell ergaben sich folgende Indizes: $\chi^2 = 26.39$ ($df = 26$), $p = .44$, Bollen-Stine- $p = .94$, $\chi^2/df = .73$, CFI = .99, TLI = .99, SRMR = .0421, RMSEA = .008, CI = [.000;.054], $p_{close} = .93$. Das Messmodell kann somit auch als parallel gelten. In Tabelle 50 sind die Reliabilitätskoeffizienten unter den verschiedenen Äquivalenzbedingungen dargestellt.

Tabelle 50: Reliabilitäten des Faktors Autoritäres Wissen

Reliabilitäten
Skalenreliabilität ρ kongenerisch = .72
Skalenreliabilität ρ tau = .72
Skalenreliabilität ρ parallel = .72
Cronbach's $\alpha = .72$

Die Skala enthält Items, die sich auf die Aspekte der Wissensvermittlung durch Autoritäten beziehen, sowie Items, welche anzeigen, dass eine fortgesetzte Beschäftigung mit offenen Fragen nicht lohnenswert ist bzw. sich nur für besonders begabte Menschen auszahlt. Ein

Item thematisiert auch die Schnelligkeit des Lernens, was aber auch im Hinblick auf Begabung gesehen werden kann, zwei weitere Items beziehen sich darauf, dass neu Gelerntes nicht mit bisher erworbenem Wissen in Beziehung gesetzt werden muss. Die Skala wird deswegen im Folgenden als *Autoritäres-einfaches Wissen* bezeichnet.

4.1.4.2 Skala F2: Existenz von Wahrheit

Der zweite Faktor besteht aus insgesamt sieben Items (Nr. 21, 12, 33, 6, 59, 34, 55). Keine der Variablen wies eine merkliche Abweichung von der Normalverteilung auf, alle Schiefen und Wölbungen sind in den Grenzen von West et al. (1995). Das Modell zeigte eine akzeptable Modellpassung ($\chi^2 = 32.08$ ($df = 14$), $p = .004$, Bollen-Stine- $p = .03$, $\chi^2/df = 2.29$, CFI = .90, TLI = .86, SRMR = .0529, RMSEA = .076, CI = [.041;.111], $p_{close} = .10$). Die Modifikationsindizes ergaben aber einen Hinweis auf einen möglicherweise korrelierenden Messfehler zwischen Item 12 und 21. Beide Items beinhalten die Aussage, dass durch Wissenschaftler wahre Aussagen gefunden werden können, zudem beinhalten beide Items das Wort Wahrheit. Ein korrelierender Messfehler deutet also auf einen Methodeneffekt bei diesen Items hin. Beide Items weisen eine ähnliche hohe Faktorladung auf. Da die Formulierung von Item 12 weitaus komplizierter ist als die von Item 21 und dadurch leichter von Versuchspersonen unterschiedlich interpretiert werden kann, wird Item 12 aus einer weiteren psychometrischen Analyse ausgeschlossen.

Eine erneute Schätzung des Modells ohne Item 12 ergab eine gute Modellpassung ($\chi^2 = 8.62$ ($df = 9$), $p = .47$, Bollen-Stine- $p = .63$, $\chi^2/df = .96$, CFI = 1.00, TLI = 1.00, SRMR = .0338, RMSEA = .000, CI = [.073;.111], $p_{close} = .81$). Der Ausschluss von Item 12 brachte eine deutliche Verbesserung der Modellpassung. Damit kann man die Items als eindimensional ansehen, sodass auch ein kongenerisches Messmodell vorliegt, siehe siehe Tabelle 51. Es folgt die Überprüfung, ob das Messmodell die Bedingungen der τ -Äquivalenz bzw. Parallelität erfüllt.

Tabelle 51: Faktorladungen des Faktors Existenz von Wahrheit

Item	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
Item 21	1.00	.45			
Item 33	.98	.47	.23	4.20	***
Item 06	1.21	.58	.26	4.63	***
Item 59	1.12	.57	.24	4.59	***
Item 34	1.32	.45	.32	4.14	***
Item 55	.89	.45	.22	4.11	***

Anmerkung: ***: $p < .001$

Die Überprüfung der Restriktionen des τ -äquivalenten Messmodell zeigte eine gute Modellpassung ($\chi^2 = 12.58$ ($df = 14$), $p = .56$, Bollen-Stine- $p = .71$, $\chi^2/df = .89$, CFI = 1.00, TLI = 1.01, SRMR = .0447, RMSEA = .000, CI = [.000;.059], $p_{close} = .90$).

Das parallele Messmodell wies hingegen keine akzeptable Modellpassung mehr auf ($\chi^2 = 73.04$ ($df = 19$), $p < .001$, Bollen-Stine- $p = .001$, $\chi^2/df = 3.84$, CFI = .58, TLI = .67, SRMR = .0452, RMSEA = .113, CI = [.086;.141], $p_{close} < .01$). somit kann das Messmodell als τ -Äquivalent, aber nicht als parallel angenommen werden. In Tabelle 52 sind die Reliabilitätskoeffizienten für die Skala ausgeführt.

Tabelle 52: Reliabilitäten der Skala Existenz von Wahrheit

Reliabilitäten
Skalenreliabilität ρ kongenerisch = .65
Skalenreliabilität ρ tau = .64
Cronbach's $\alpha = .65$

Die Skala enthält Items, die sich auf die Existenz bzw. Erkennbarkeit und Dauerhaftigkeit von Wahrheit beziehen. Ein Item bezieht sich darauf, dass fast ausschließlich glaubwürdige und damit auch wahre Fakten vermittelt werden. Item 33 weicht vom Inhalt der übrigen Items ab, stellt aber mit dem Bezug auf eine eindeutige Bedeutung von Wörtern auch einen Bezug zum Wahrheitsgehalt von Aussagen her. Das Item 55 bezieht sich auf die Veränderbarkeit von Begabung und fällt somit aus dem Bezugsrahmen der übrigen Items heraus. Da sich die überwiegende Mehrzahl der Items auf Aspekte der Wahrheit und deren Entdeckbarkeit abzielt, wird diese Skala im Folgenden *Existenz von Wahrheit* genannt werden.

4.1.4.3 Skala F3: Integratives Wissen

Der dritte Faktor wird aus den Items 61, 35, 10, 18, 9 und 54 gebildet. Keines der Items wies eine merkliche Abweichung von der Normalverteilung auf, alle Schiefen und Wölbungen sind in den Grenzen von West et al. (1995). Die Fitindizes des Messmodells waren $\chi^2 = 10.66$ ($df = 9$), $p = .30$, Bollen-Stine- $p = .43$, $\chi^2/df = 1.18$, CFI = .99, TLI = .98, SRMR = .0369 und RMSEA = .029, CI = [.000;.084], $p_{close} = .67$. Somit liegt eine gute Modellpassung vor, sodass keine weitere Veränderung an der Skala notwendig ist und ein kongenerisches Messmodell vorliegt.

Tabelle 53: Faktorladungen des Faktors Integratives Wissen)

Item	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
Item 61	1.00	.54			
Item 35	1.08	.62	.21	5.26	***
Item 10	.82	.42	.19	4.31	***
Item 18	1.17	.53	.24	4.96	***
Item 09	.68	.33	.19	3.62	***
Item 54	.79	.48	.17	4.66	***

Anmerkung: ***: $p < .001$

Für das Modell mit den Restriktionen der τ -Äquivalenz liegt eine gute Modellpassung vor ($\chi^2 = 17.47$ ($df = 14$), $p = .23$, Bollen-Stine- $p = .39$, $\chi^2/df = 1.25$, CFI = .97, TLI = .97, SRMR = .0527, RMSEA = .033, CI = [.000;.077], $p_{close} = .69$).

Die Fitindizes für das parallele Modell sind: $\chi^2 = 48.97$ ($df = 19$), $p < .001$, Bollen-Stine- $p = 0.01$, $\chi^2/df = 2.58$, CFI = .77, TLI = .82, SRMR = .0596 und RMSEA = .084, CI = [.056;.114], $p_{close} = .03$. Da der CFI-Wert erheblich unter dem erforderlichen Wert für eine akzeptable Modellpassung liegt, wird nicht von einem parallelen Messmodell ausgegangen. Die Reliabilitätsschätzungen sind in Tabelle 54 aufgeführt.

Tabelle 54: Reliabilitäten der Skala Integratives Wissen

Reliabilitäten
Skalenreliabilität ρ kongenerisch = .65
Skalenreliabilität ρ tau = .64
Cronbach's $\alpha = .64$

Inhaltlich behandeln die meisten Items den Sachverhalt, dass neu erworbenes Wissen dem eigenen Vorwissen entsprechend geordnet werden sollte und nicht aus isolierten Fakten, sondern aus den Zusammenhängen übergeordneter Ideen besteht. Item 09 fällt aus diesem Schema heraus, da es sich auf Charakteristika von Lehrenden bei der Wissensvermittlung bezieht. Allerdings kann dieses Item auch in der Hinsicht interpretiert werden, dass es eine wünschenswerte Eigenschaft von Dozenten wäre, zu der jeweils referierten Thematik die zentralen Inhalte klar zu kommunizieren. Item 10 lässt sich hingegen nur schwer im Sinn der Strukturierung von Wissen als Sammlung grundlegender Konzepte verstehen, da es direkt auf die Begabung von Lernenden abzielt. Dieses Item weist auch die geringste Faktorladung auf. Da keine Anhaltspunkte vorliegen, welche die Entfernung des Items aus der Skala erforderlich

machen, wird es in der Skala belassen. Fasst man den Inhalt dieser Items zusammen, so lässt sich diese Skala als *Integratives Wissen* bezeichnen.

4.1.4.4 Faktor 4: Lernmethodik

Das Messmodell für diesen Faktor besteht aus den Items 24, 43, 25, 39 und 32. Keine der Variablen zeigte eine merkliche Abweichung von der Normalverteilung, alle Schiefen und Wölbungen sind in den Grenzen von West et al. (1995). Die Fitindizes des Messmodells deuten fast alle Fitindizes auf eine akzeptable Modellpassung ($\chi^2 = 13.63$ ($df = 5$), $p = .02$, Bollen-Stine- $p = .06$, $\chi^2/df = 2.73$, CFI = .93, TLI = .86, SRMR = .0473, RMSEA = .088, CI = [.033;.146], $p_{close} = .10$). Die Modifikationsindizes und die standardisierten Residuen legten keine weitere Modellmodifikation nahe.

Eine Inspektion der Items ergab, dass sie sich mehrheitlich auf Methoden des Wissenserwerbs beziehen. Das Item 25 ist nicht eindeutig in seinem Bezug auf die Person des Beurteilers. Das Item kann im Sinn einer Selbst- oder Fremdbeurteilung interpretiert werden. Aus diesem Grund wurde das Item aus der Skala entfernt und die Modellpassung des Messmodells erneut geprüft, es ergaben sich folgende Anpassungsindizes: $\chi^2 = 1.39$ ($df = 2$), $p = .50$, Bollen-Stine- $p = .54$, $\chi^2/df = .69$, CFI = 1.00, TLI = 1.02, SRMR = .0180 und RMSEA = .000, CI = [.000;.119], $p_{close} = .66$. Der Ausschluss von Item 25 führte also dazu, dass alle Fitindizes im Bereich einer guten Modellpassung liegen. Die Parameter des Modells sind in Tabelle 55 angegeben.

Tabelle 55: Faktorladungen des Faktor Lernmethodik

Item	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
Item 24	1.00	.52			
Item 43	1.15	.61	.26	4.43	***
Item 39	.99	.54	.23	4.39	***
Item 32	.91	.42	.23	3.92	***

Anmerkung: ***: $p < .001$

Als Nächstes wird geprüft, ob es sich um ein τ -äquivalentes oder paralleles Messmodell handelt. Alle Indizes liegen im Bereich guter Modellpassung, es liegt ein τ -äquivalentes Messmodell vor ($\chi^2 = 2.36$ ($df = 5$), $p = .80$, Bollen-Stine- $p = .90$, $\chi^2/df = .47$, CFI = 1.00, TLI = 1.04, SRMR = .0246, RMSEA = .000, CI = [.000;.060], $p_{close} = .92$).

Für das Messmodell mit den notwendigen Restriktionen für Parallelität liegen alle Indizes im Bereich einer guten Modellpassung, es liegt also auch ein paralleles Messmodell vor ($\chi^2 = 12.22$ ($df = 8$), $p = .14$, Bollen-Stine- $p = .40$, $\chi^2/df = 1.57$, CFI = .96, TLI = .96, SRMR = .0350, RMSEA = .049, CI = [.000;.100], $p_{close} = .46$).

Tabelle 56: Reliabilitäten der Skala Lernmethodik

Reliabilitäten
Skalenreliabilität ρ kongenerisch = .60
Skalenreliabilität ρ tau = .60
Skalenreliabilität ρ parallel = .59
Cronbachs $\alpha = .60$

Die inhaltliche Interpretation dieser Skala gestaltet sich geradlinig. Alle Items zielen auf das Lernverhalten ab und beziehen sich auf Methoden und Vorgehensweisen zum Wissenserwerb. Im Folgenden wird diese Skala deshalb *Lernmethodik* genannt.

4.1.4.5 Faktor 5: Unsicherheit des Wissens

Das anfängliche Messmodell für Faktor 5 bestand aus den Items 27, 15, 14, 46, 42, 48 und 30. Keines der Items wies eine merkliche Abweichung von der Normalverteilung auf, alle Schiefen und Wölbungen sind in den Grenzen von West et al. (1995). Die Fitindizes für dieses Modell waren $\chi^2 = 38.45$ ($df = 14$), $p < .000$, Bollen-Stine- $p = .002$, $\chi^2/df = 2.75$, CFI = .79, TLI = .69, SRMR = .0665 und RMSEA = .089, CI = [.056;.123], $p_{close} = .03$. Der CFI- und TLI-Wert liegt weit unterhalb der Grenze einer akzeptablen Modellpassung. Eine Inspektion der Modifikationsindizes ergab, dass es zwischen Item 48 und 30 einen korrelierenden Messfehler gibt. Inhaltlich beziehen sich die beiden Items auf den Aspekt, dass es nichts Sicheres gibt bzw. die Ungewissheit ein zentrales Merkmal der Existenz ist. Da das Item 48 sich aber nicht direkt auf Wissen bzw. Wissenserwerb bezieht, sondern auf alle erdenklichen Situationen, welche mit Unsicherheit behaftet sind, wird dieses Item aus der Skala ausgeschlossen.

Die erneute Untersuchung ohne Item 48 ergab folgende Fitindizes: $\chi^2 = 16.22$ ($df = 9$), $p = .06$, Bollen-Stine- $p = .10$, $\chi^2/df = 1.80$, CFI = .92, TLI = .87, SRMR = .0473 und RMSEA = .060 (CI = [.000;.106], $p_{close} = .32$). Die Fitindizes sind fast alle im Bereich einer guten bis akzeptablen Modellpassung. Die Modifikationsindizes deuteten aber auf eine Fehlspezifikation der Beziehung zwischen den Items 14 und 30, wobei das standardisierte Residuum von -1.99 auf eine Unterschätzung der tatsächlichen Korrelation und der vom Modell implizierten Korrelation hindeutet. Da beide Items inhaltlich nicht ähnlich sind, liegt es nahe, keine korre-

lierenden Messfehler als Grund für diese Unterschätzung anzunehmen. Item 14 bezieht sich auf die Integration verschiedener Wissensbereiche und ist somit semantisch den anderen Items dieses Faktors nicht ähnlich, es wird deswegen aus der Skala ausgeschlossen.

Eine erneute Berechnung des Modells weist eine gute Modellpassung auf ($\chi^2 = 0.93$ ($df = 5$), $p = .96$, Bollen-Stine- $p = .97$, $\chi^2/df = .19$, CFI = 1.00, TLI = 1.17, SRMR = .0151, RMSEA = .000 (CI = [.000;.002], $p_{close} = 0.99$). Somit liegt auch ein kongenerisches Messmodell für die Erfassung des Faktors vor. Die Faktorladungen sind in Tabelle 57 zu finden.

Tabelle 57: Faktorladungen des Faktors Unsicherheit des Wissens

Item	partielles Regressionsgewicht		SE	CR	<i>p</i>
	nicht standardisiert	standardisiert			
Item 27	1.00	.58			
Item 15	.51	.35	.16	3.10	***
Item 46	.59	.49	.17	3.53	***
Item 42u	.47	.21	.22	2.15	.03
Item 30	.80	.42	.23	3.41	***

Anmerkung: ***: $p < .001$

Nun bleibt noch zu überprüfen, ob das Messmodell die Bedingungen der τ -Äquivalenz und Parallelität erfüllen. Die Fitindizes für das Messmodell mit den Restriktionen der τ -Äquivalenz legen eine gute Modellpassung nahe, sodass von einem τ -äquivalenten Messmodell ausgegangen werden kann ($\chi^2 = 7.87$ ($df = 9$), $p = .55$, Bollen-Stine- $p = .58$, $\chi^2/df = .88$, CFI = 1.00, TLI = 1.03, SRMR = .046, RMSEA = .000, CI = [.000;.069], $p_{close} = .85$).

Die Fitindizes für das parallele Messmodell waren: $\chi^2 = 116.78$ ($df = 13$), $p < .01$, Bollen-Stine- $p = 0.001$, $\chi^2/df = 8.98$, CFI = .00, TLI = .668, SRMR = .0568 und RMSEA = .190, CI = [.159;.222], $p_{close} < .01$. Es liegt somit kein paralleles Messmodell vor. Die Reliabilitäts-schätzungen sind in Tabelle 58 zu finden.

Tabelle 58: Reliabilitäten für die Skala Unsicherheit des Wissens

Reliabilitäten
Skalenreliabilität ρ kongenerisch = .47
Skalenreliabilität ρ tau = .46
Cronbachs $\alpha = .46$

Die Interpretation dieser Skala gestaltet sich schwieriger als bei den anderen Skalen, da sich alle Items inhaltlich unterscheiden. Die Items 46, 42 und 30 zielen auf einen Unsicherheits-

faktor ab, wobei nur das Item 46 einen mehr oder weniger direkten Bezug zum Wissen aufweist, Item 42 bezieht sich auf eine Präferenz für Filme mit offenem Ende und Item 30 bezieht sich auf die Unsicherheit, die dem Leben allgemein zugeschrieben wird. Item 27 hingegen bezieht sich nur indirekt auf die Unsicherheit von Wissen, es zielt mehr auf die Präferenz des Nachdenkens über schwierige Frage ab. Item 15 hingegen hat einen völlig anderen Bezug auf die Erkennbarkeit der Lernfähigkeit. Da sich die Mehrzahl der Items auf Komponenten der Unsicherheit bezieht, wird diese Skala im Folgenden als *Unsicherheit des Wissens* bezeichnet.

4.2 Diskussion der Analysen des EBQ

4.2.1 Das Modell von Schommer

Im ersten Schritt der Analyse des Fragebogens wurde die Zuordnung der Items zu den Subskalen, wie sie von Schommer (1990) vorgegeben wurden, mithilfe einer CFA geprüft. Im zweiten Schritt wurde die Zuordnung der Subskalen zu den Faktoren geprüft. Überdies wurde mittels einer EFA der Subskalen die faktorielle Struktur ermittelt und mit den Ergebnissen von Schommer (1990, 1992, 1993a) verglichen.

Die Überprüfung der Zuordnung der Items zu den Subskalen mithilfe von eindimensionalen Faktormodellen führte bis auf die Skalen S6 und S10 zu Modellen, die keinen akzeptablen Modellfit aufwiesen. Dieses Ergebnis legt nahe, dass diese Messmodelle falsch spezifiziert sind, z. B. können Items vorliegen, die nicht durch einen gemeinsamen Faktor im Sinn des *Common-factor*-Modells von Thurstone (1947) beeinflusst werden, oder es können Methodeneffekte wie korrelierte Messfehler aufgetreten sein. Bei den Modellen für die Skalen S1 und S5 traten auch Fälle der empirischen Unteridentifikation auf, bei denen das Modell nicht geschätzt werden kann (vgl. Brown, 2006). Inhaltlich bedeutet der fehlende Zusammenhang der manifesten Variablen, dass diese Items keinen gemeinsamen Faktor aufweisen. Somit erweist sich auch die vorgenommene Zuordnung der Items zu den Subskalen durch Schommer bis auf die Skalen S6 und S10 als nicht haltbar.

Trotz der nicht passenden Messmodelle wurden zur Prüfung des Gesamtmodells die Items zu vorgegebenen Summenskalen addiert und anschließend die Zuordnung dieser Summenskalen zu den fünf postulierten epistemologischen Dimensionen mithilfe einer konfirmatorischen Faktorenanalyse überprüft. Dieses 5-faktorielle Modell wies einen signifikanten χ^2 -Wert sowie einen CFI und TLI auf, der unterhalb des für einen ausreichenden Modellfit notwendigen Bereichs lag. Somit kann auch hier davon ausgegangen werden, dass die vorgegebene Zuordnung von Skalen zu Faktoren nicht haltbar ist und das Modell mit fünf Faktoren auf diese Weise also nicht bestätigt werden kann. Allerdings ist bei der Prüfung des Gesamtmodells und der Interpretation der Ergebnisse die Verwendung der Itempakete, die von Schommer vorgegeben wurden (vgl. Kapitel 1.3.1 sowie Anhang A zur genauen Bildung der Itempakete), kritisch zu betrachten. Die Bildung von Itempaketen bietet Vorteile, weist aber auch entscheidende methodische Nachteile auf. Nach Cole et al. (2000) widerspricht dieses Vorgehen der eigentlichen Intention der Faktorenanalyse, den Items zugrunde liegende Dimensionen zu

entdecken. Die Verwendung von Itempaketen „versteckt“ die Beziehungen der Items untereinander. Bei der Bildung der Itempakete besteht überdies die Möglichkeit zur Zusammenfassung von Items, die nicht zu dem gleichen Konstrukt gehören, also multidimensional statt eindimensional sind. In diesem Fall kann weder die Quelle der Varianz der latenten Variablen eindeutig festgestellt werden, noch kann eine eindeutige Interpretation der latenten Variable erfolgen. Auch werden bei Multidimensionalität der zusammengefassten Items die Beziehungen der latenten Variablen falsch geschätzt (Little, Cunningham, Shahar & Widaman, 2002, Bandalos, 2002). Zudem kann die Verwendung von Itempaketen dazu führen, dass mögliche Fehlspezifikationen nicht erkannt werden, im Gegensatz zu einer Prüfung der Modellstruktur auf Basis der Items (Little et al., 2002). Bandalos und Finney (2001) weisen überdies darauf hin, dass bei der Verwendung von Itempaketen auch automatisch die Passung des Modells verbessert, selbst wenn eine Fehlspezifikation vorliegt. Dies spricht somit dafür, dass die Passung des Modells somit noch schlechter ist, als dies durch die nicht akzeptable Modellpassung nahegelegt wird.

Um die vorliegenden Ergebnisse weiterhin mit den Ergebnissen der Untersuchungen von Schommer (1990, 1992, 1993a) vergleichen zu können, wurden die Subskalen zusätzlich noch einer explorativen Faktorenanalyse unterzogen. Die Ergebnisse werden in Tabelle 41 mit den Faktorenstrukturen aus Schommer (1990, 1992, 1993a) verglichen. Es wurden auf diese Weise vier Faktoren extrahiert und einer Varimax-Rotation unterzogen, wobei das KMO-Maß aber nur .68 betrug. Die aufgeklärte Varianz betrug auch nur 31.2 %. Verglichen mit 55.2 % bei Schommer (1990) und 54.2 % in Schommer et al. (1992) sind dies im Schnitt ca. 23 % Varianz, die weniger aufgeklärt wird. In der ebenfalls als direkte Replikation von Schommer ausgelegten Studie von Clarebout et al. (2001) wurden in zwei unterschiedlichen Stichproben jeweils eine Varianzaufklärung von 54.84 % bzw. 50.31 % erreicht. Ein möglicher Grund für die doch eher geringe Varianzaufklärung ist, dass die Subskalen selbst eine eher heterogene Itemzusammenstellung darstellen, was auch durch die überwiegend nicht passenden Messmodelle der Subskalen demonstriert wird. Zudem teilen die Subskalen an sich nur wenig Varianz, was durch das niedrige KMO-Maß nahegelegt wird.

Es zeigte sich jedoch, dass jede Subskala, die in der hier durchgeführten Untersuchung eine Faktorladung größer als .40 aufwies, einen vergleichbaren Wert auch in den genannten Studien von Schommer erzielte (vgl. Tabelle 41). Allerdings ist im Vergleich mit der Faktorenstruktur zu Schommer keine Gemeinsamkeit zu erkennen. Diese merklichen Unterschiede in

dem Ladungsmuster sind ebenso bei Clarebout et al. (2001) zu finden, die in ihren beiden Studien ebenso Ladungsmuster erhielten, die dem von Schommer widersprechen.

Die 4-faktorielle Lösung von Schommer selbst ist aber ebenso fraglich. In den Untersuchungen von 1992 und 1993 wurde das Eigenwert-Kriterium von Kaiser angepasst, um zu einer 4-faktoriellen Lösung zu gelangen, es wurde – wenn auch nur minimal – der Eigenwert herabgesetzt, um diese Faktorenlösung zu rechtfertigen. Bei Schommer (1992) wurde die 4-faktorielle Lösung gewählt, weil diese bei einer konfirmatorischen Faktorenanalyse die bessere Modellpassung ergab. Allerdings gelten auch für diesen Befund die bereits vorgebrachten Argumente gegen eine Verwendung von Itempaketen bei einer CFA, die sich auf die Verwendung von Itempaketen bei der Berechnung einer EFA übertragen lassen.

Somit kann die erste Forschungsfrage zu den Modellvorstellungen von Schommer beantwortet werden. Die postulierten Skalen sind durch eine konfirmatorische Faktorenanalyse nicht zu bestätigen, ebenso lassen sich die fünf epistemologischen Dimensionen weder durch eine konfirmatorische noch durch eine explorative Faktorenanalyse bestätigen. Überdies erscheint die Verwendung der Itempakete aus methodischer Sicht fragwürdig.

4.2.2 Auswertung des EBQ auf Itemebene

Die zweite Forschungsfrage zielt auf die Struktur des EBQ bei einer Analyse auf Itemebene ab, d. h. welche Faktorenstruktur sich ergibt, wenn auf Basis der Items faktorisiert wird. Die Auswertung des Schommer-Fragebogens auf Itemebene führte zu einer 5-faktoriellen Faktorenlösung. Ein erstes Problem bei der faktorenanalytischen Auswertung stellt die schlechte Eignung zur Durchführung einer Faktorenanalyse dar, die durch das mäßige KMO-Maß nahegelegt wurde. Das KMO-Maß stellt zwar kein inferenzstatistisches Kriterium zur Prüfung der Dateneignung dar, wird aber in der Literatur zur Faktorenanalyse als eines der besten Kriterien angesehen (Stewart, 1981). Ein geringes KMO-Maß deutet darauf hin, dass die Items jeweils nur einen geringen Anteil an gemeinsamer Varianz haben. Von der Interpretation ähnlich sind die MSA-Werte. Ein geringer MSA-Wert deutet darauf hin, dass ein Item nur einen geringen Teil seiner Varianz mit den anderen Items teilt. Üblicherweise wird ein KMO- bzw. MSA-Wert größer als .80 als mit der Durchführung einer Faktorenanalyse vereinbar angesehen (Backhaus et al., 2003).

In der durchgeführten Analyse wurde ein wesentlich liberaleres Kriterium verwendet, und zwar sollte der MSA-Wert größer als .60 sein. Die Wahl eines liberalen MSA-Kriteriums hatte in erster Linie pragmatische Gründe und schloss sich der Empfehlung von Bühner (2006) an, der als Richtwert ein MSA-Wert größer als .60 empfiehlt. Keiner der MSA-Werte der Items vor der Ausschlussprozedur war größer als .80, sodass eine Datenanalyse bei Verwendung eines strengeren Kriteriums nicht möglich gewesen wäre. Das niedrige KMO-Maß von .64 deutet darauf hin, dass die Items jeweils nur einen geringen Anteil an gemeinsamer Varianz haben und damit einhergehend eher gering korrelieren. Dies bedeutet aber auch, dass die Versuchspersonen nur wenig Ähnlichkeit zwischen den Items erkennen. Durch den Ausschluss derjenigen Items, die MSA-Werte unter .60 aufwiesen, konnte das KMO-Maß auf .73 angehoben werden. Die Versuchspersonen nahmen nach dem Ausschluss der Items mit geringem KMO-Maß also die verbleibenden Items als ähnlicher wahr, wobei diese Zunahme selbst wieder nur gering ist. Die Items teilen auch nach der Ausschlussprozedur nur wenig Varianz.

Eine Inspektion der ausgeschlossenen Items zeigt, dass sich diese auf inhaltlich sehr unterschiedliche Sachverhalte beziehen. Beispiele für ausgeschlossene Items sind: „Ein Satz hat wenig Bedeutung, solange man den Kontext, in dem er gesprochen oder geschrieben wurde, nicht kennt“, „Wie viel mal lernt, hängt in erster Linie vom Dozenten ab“, „Die Dozenten machen Dinge komplizierter als sie sind“, „Die wichtigsten Aspekte wissenschaftlichen Arbeitens sind originelle Ideen“, „Ich verstehe schwierige Sachverhalte, wenn ich alle Störungen von außen abschalte und mich wirklich konzentriere“ und „Ein geordneter Kopf ist ein leerer Kopf“. Wie die Beispiele verdeutlichen, weisen die Items inhaltlich keinen Bezug zueinander auf und zum anderen ist unklar, in welcher Weise die Items für die Erfassung epistemologischer Überzeugungen relevant sind. Berechnet man das KMO-Maß für die Itemmenge, die während der Selektion ausgeschlossen wurde, so ergibt sich .52. Dies deutet darauf hin, dass die ausgeschlossenen Items an sich wenig Varianz gemeinsam haben, also von den Versuchspersonen nicht als ähnlich beurteilt wurden.

Das KMO-Maß von .73 für die verbleibenden Items nach der Ausschlussprozedur ist überdies immer noch unter dem von Backhaus et al. (2003) geforderten Wert von .80, der als notwendig für die Durchführung einer EFA angesehen wird. Auch nach dem Ausschluss der Items mit niedrigen MSA-Werten teilen die Items anscheinend nicht ausreichend Varianz, um ein entsprechend hohes KMO-Maß zu erreichen. Allerdings waren vor dem Ausschlussverfahren keine Items mit MSA-Werten größer als .80 vorhanden, nach den zwei Ausschlussritten

waren es drei Items (vgl. Tabelle 42). Aus diesem Grunde wurde entschieden, die Faktorenanalyse trotz des ungeeigneten KMO-Maßes durchzuführen. Von ursprünglich 63 Items blieben nach dem Ausschlussverfahren noch 38 Items übrig. Eine ähnlich große Zahl von Items, die aus der weiteren Auswertung ausgeschlossen wurden, ist auch in den Untersuchungen zum EBQ von Bendixen et al. (1994) und Schommer-Aikins, Brookhart, Hutter und Mau (2000) vorhanden.

Im Hinblick auf die Bestimmung der Faktorenanzahl existieren mehrere verschiedene Methoden, hier wurden insbesondere die Parallelanalyse und der Scree-Plot betrachtet (vgl. Kapitel 3.1). Grundbedingung war die inhaltlich adäquate Interpretierbarkeit der gefundenen Lösung. Dies führte zur Beibehaltung der 5-faktoriellen Lösung. Die Faktorzahl ist vergleichbar mit anderen Studien, die ebenfalls auf Basis der Items faktorisierten. So fanden auch Jehng et al. (1993) mit ihrer Weiterentwicklung des EBQ eine 5-faktorielle Lösung, ebenso Wood und Kardash (2002), die ebenfalls den Fragebogen von Jehng et al. (1993) verwendeten. Eine Studie, die den Fragebogen von Schommer direkt auf Itemebene faktorisierte, ist die von Schraw et al. (2002). Auch diese Autoren fanden unter der Verwendung von Hauptachsenanalysen mit obliquen Rotationen eine 5-faktorielle Lösung. Allerdings basiert deren 5-faktorielle Lösung auf der Berechnung einer Lösung mit 19 Faktoren, von denen anschließend nur die ersten fünf Faktoren betrachtet wurden.

Um die 5-faktorielle Faktorenlösung weiterhin auf ihre Angemessenheit zu prüfen, wurde ein E/CFA-Modell aufgestellt. Das aufgestellte E/CFA-Modell zeigte eine gute Modellanpassung, allerdings ergab sich ein signifikanter χ^2 -Test sowie ein CFI-Wert, der unterhalb der akzeptablen Schwelle lag. Der signifikante χ^2 -Test deutet auf eine mangelnde Passung des Modells mit den Daten hin. Diese kann im Zusammenhang mit nicht spezifizierten Fehlerkovarianzen stehen. Hinweise auf Fehlspezifikationen geben die Modifikationsindizes, die für einige Fehlerkovarianzen eine Verbesserung des χ^2 -Wertes andeuteten. An dieser Stelle wurde allerdings auf eine Modellierung der Fehlerkovarianzen zur Verbesserung der Modellpassung des E/CFA-Modells verzichtet, da im Anschluss an die faktorenanalytische Auswertung noch eine Skalenkonstruktion erfolgen sollte, bei welcher korrelierende Fehlerterme separat untersucht werden. Der geringe Wert des CFI könnte durch die Komplexität des E/CFA-Modells bedingt sein, da der CFI komplexe Modelle abstrahiert (Schermelleh-Engel et al., 2003). Ein E/CFA-Modell ist in der Hinsicht komplex, dass alle möglichen Kreuzladungen modelliert werden (mit Ausnahme der Skalierer, vgl. Brown, 2006). Das E/CFA-Modell spricht somit auch für

die vorgestellte 5-Faktorenlösung. Das aufgestellte Faktormodell mit fünf Faktoren scheint somit die Daten angemessen zu beschreiben.

Um den Zusammenhang der Faktoren zu ermitteln, wurden die Korrelationen der latenten Variablen geschätzt. Schommer (1990) hat ihre Dimensionen als „mehr oder weniger“ unabhängig charakterisiert. Um die Frage nach der Zusammenhangstruktur zu beantworten, wurde dem Vorschlag von Bühner (2006) gefolgt und bei der EFA zuerst die oblique Promax-Rotation gewählt. Da die Faktoren deutliche Korrelationen zeigten, wurde die oblique Rotation beibehalten. Ein Nachteil weist die Promax-Rotation auf, die Wahl des Parameters κ bestimmt die mögliche Höhe der Korrelation der Faktoren. Die gefundenen Korrelationen zwischen den Faktoren bilden somit nicht die wirklichen Korrelationen der Konstrukte ab, sondern nur die Korrelationen, wie sie bei der Wahl von κ möglich sind. Dieser Nachteil findet sich aber nicht bei E/CFA-Modellen, da hier die Korrelationen zwischen den latenten Variablen frei geschätzt werden und es nicht notwendig ist, eine Rotationstechnik zu verwenden (Brown, 2006). In Tabelle 48 sind die Korrelationen aus der EFA und dem E/CFA-Modell gegenübergestellt. Teilweise ergeben sich bei beiden Verfahren andere Werte. Trotz dieser Unterschiede der numerischen Werte der Korrelationen zeigen sich deutliche Abhängigkeiten zwischen den ermittelten Dimensionen epistemologischer Überzeugungen, somit können diese also nicht als orthogonal angenommen werden.

Damit kann die Forschungsfrage nach der faktoriellen Struktur des Schommer-Fragebogens beantwortet werden. Die Items weisen eine 5-faktorielle Struktur auf. Allerdings kann dieses Ergebnis nicht ohne Weiteres mit anderen Analysen des Schommer-Fragebogens verglichen werden, da diese entweder nicht auf Itemebene durchgeführt wurden, ein anderes Itemset verwendeten oder eine andere Methode bei der Faktorenextraktion angewandt wurde. Die Faktoren zeigten sowohl bei der Berechnung mit einer Promax-Rotation im Kontext der EFA als auch bei der Berechnung der Korrelationen latenter Variablen im Kontext der E/CFA eine deutliche Korrelation. Im Sinne der Konzeption von Schommer (1990) sind die Faktoren somit eher abhängig als unabhängig.

4.2.3 Abgeleitete Skalen und deren psychometrische Eigenschaften

Die inhaltliche Deutung der Faktoren wird im Folgenden zusammen mit der Konstruktion von Skalen, die zur Messung der gefundenen Faktoren gebraucht werden können, und deren psychometrischen Eigenschaften besprochen. Die erste Skala wurde mit Bezeichnung *Autori-*

täres-einfaches Wissen versehen. Sie beinhaltet Items, die sich auf verschiedene Wissensaspekte beziehen. Ein eindimensionales CFA-Modell wies einen guten Modellfit hin, ebenso die Modelle mit den Restriktionen für τ -Äquivalenz und Parallelität. Die Skalenreliabilität und die Schätzung der Reliabilität durch Cronbachs α sind wegen der Parallelität auch identisch mit $\alpha = .72$ und $\rho = .72$. Allerdings ist die Reliabilität am unteren Ende des Bereichs, der üblicherweise als akzeptabel angesehen wird (Weise, 1973). Die Benennung wurde hauptsächlich wegen der beiden Items 31, 13 und 50 gewählt (vgl. Tabelle 44). Die ersten beiden Items beziehen sich direkt auf den Aspekt von Autoritäten bei der Wissensvermittlung, während sich das Item 50 indirekt auf Autorität bezieht, in dem es die langfristige Beschäftigung mit Wissensinhalten nur besonders begabten Studenten zuschreibt. Die übrigen Items beziehen sich eher auf den Aspekt, dass man schnell lernt und das Gelernte nicht mit dem bisher erworbenen Wissen in Verbindung bringen muss. Diese unterschiedlichen Aspekte im Umgang mit Wissen, auf die sich die Items beziehen, stellen die Homogenität der Items infrage, wobei hier gemeint ist, dass sich der Faktor evtl. in verschiedene Facetten aufteilt. Einen Hinweis hierauf kann die Präzision von Cronbachs α geben (Cortina, 1993). Dieser Kennwert gibt zusammen mit der mittleren Item-Interkorrelation Hinweise auf eine mögliche Multidimensionalität einer Skala. Die mittlere Item-Interkorrelation sollte zwischen .20 und .40 liegen (Briggs & Cheek, 1993), die Präzision von Cronbachs α sollte .01 nicht übersteigen. Werte größer als .01 können schon auf Multidimensionalität hinweisen (Cortina, 1993). Die mittlere Item-Interkorrelation der Skala für die Messung des ersten Faktors *Autoritäres-einfaches Wissen* liegt bei .27 im akzeptablen Bereich, aber die Präzision von Cronbachs α ist mit .01 grenzwertig (s. Tabelle 59). somit liegt ein Hinweis auf Multidimensionalität vor³⁹.

Tabelle 59: Mittlere Item-Interkorrelationen und Präzision von Cronbachs α

Skala	Mittlere Item-Interkorrelation	Präzision α
F1	.27	.010
F2	.24	.014
F3	.24	.018
F4	.27	.020
F5	.17	.023

Zur genaueren Untersuchung, ob sich die Skala ggf. in Facetten gliedern lässt, werden die Items ausgehend von deren Inhalt in zwei Gruppen eingeteilt. Die erste Gruppe wird aus den

³⁹ Auch wenn es nicht üblich ist bei der Diskussion empirischer Analysen neue Berechnungen hinzuzufügen, wird dies in der vorliegenden Arbeit getan, um die Diskussion auf eine empirische Grundlage zu stellen. Es sei aber darauf hingewiesen, dass die angestellten Berechnungen nur die Plausibilität der diskutierten Sachverhalte unterstützen soll und auf jeden Fall an einer unabhängigen Stichprobe repliziert werden müssen.

Items 31, 13 und 50 gebildet und deckt den Bereich ab, der sich auf den autoritären Aspekt bei der Beschäftigung mit Lernmaterial bezieht. Die zweite Gruppe bilden die Items 44, 51, 52 und 63, die sich auf die Verbindung von neuem und altem Wissen, die längerfristige Beschäftigung mit Lernmaterial sowie schnellem Lernen beziehen. Es wird hierzu ein zweifaktorielles CFA-Modell berechnet. Dieses Modell zeigte einen guten Modellfit ($\chi^2 = 8.35$ ($df = 13$), $p = .82$, Bollen-Stine- $p = .99$, $\chi^2/df = 0.64$, CFI = 1.00, TLI = 1.04, SRMR = .0268, RMSEA = .000 (CI = [.000;.041], $p_{close} = .97$). Da das zweifaktorielle Modell in das einfaktorielle Modell geschachtelt⁴⁰ ist, kann mithilfe eines χ^2 -Differenztests die Verbesserung der Modellpassung durch ein zweifaktorielles Modell gegenüber dem einfaktoriellen Modell geprüft werden. Bei diesem Test wird die Differenz der χ^2 -Werte des ursprünglichen Modells und des darin geschalteten Modells berechnet, ebenso wird die Differenz der Freiheitsgrade berechnet. Mit diesen beiden Differenzwerten wird dann ein χ^2 -Test berechnet. Ist dieser signifikant, weist das geschachtelte Modell eine deutliche Verbesserung der Modellpassung auf. Für das zweifaktorielle Modell ergibt sich im vorliegenden Fall ein $\Delta\chi^2 = 0$ ($\Delta df = 1$). Das zweifaktorielle Modell bietet also keine merkliche Verbesserung der Modellpassung gegenüber dem einfaktoriellen Modell. Die identischen χ^2 -Werte erklären sich daraus, dass die Korrelation der beiden Faktoren auf 1 geschätzt wird. Da sich die Modellpassung nicht verschlechtert, kann dies als Hinweis auf eine möglicherweise vorhandene mehrfaktorielle Struktur gelten. Allerdings ist dies nicht mehr als eine begründete Vermutung, ausgehend von der Präzision von Cronbachs α , die an der oberen Grenze des akzeptablen Bereichs liegt, und der gleich guten Modellpassung des zweifaktoriellen Modells.

Die zweite Skala wurde *Existenz von Wahrheit* genannt. Die Benennung beruht auf den Items 21, 59 und 34, die eine Aussage über die Existenz von Wahrheit und die Dauerhaftigkeit von wissenschaftlichen Aussagen beinhalten. Ein eindimensionales Modell mit allen Items führte zu einem schlechten Modellfit. Die Inspektion der Residuen und Modifikationsindizes brachte Hinweise auf einen möglichen korrelierenden Messfehler zwischen Item 21 und 12, sodass Item 12 ausgeschlossen wurde. Das einfaktorielle CFA-Modell zeigte einen guten Modellfit, ebenso das Modell mit den Restriktionen für τ -Äquivalenz, allerdings ist das Messmodell nicht parallel. Bedingt durch die τ -Äquivalenz weisen auch die Skalenreliabilität ρ und Cronbachs α mit .64 bzw. .65 eine fast identische Höhe auf. Die Reliabilität liegt aber in jedem

⁴⁰ Das einfaktorielle Modell ist äquivalent zu einem zweifaktoriellen Modell, bei welchem die Korrelation der Faktoren auf eins fixiert ist (Moosbrugger & Schermelleh-Engel, 2008). Äquivalente Modelle lassen sich durch jeweils identische Kovarianzmatrizen beschreiben und sind deswegen nicht in ihrer Modellpassung zu unterscheiden (Bühner, 2006).

Fall unter der akzeptablen Grenze von .70 (Weise, 1973). Ähnlich wie bei der ersten Skala liegt zwar die mittlere Item-Interkorrelation in dem von Briggs und Cheek (1993) angegebenen Bereich, aber die Präzision von Cronbachs α gibt auch hier wiederum einen Hinweis auf eine mögliche Multidimensionalität der Skala (vgl. Tabelle 59).

Daher wurde aufgrund inhaltlicher Überlegungen – analog zu dem Vorgehen bei der ersten Skala – eine Gruppierung von Items gebildet und darauf aufbauend ein mehrdimensionales CFA-Modell angegeben. Die Items lassen sich inhaltlich in zwei Gruppen gliedern. Die erste Gruppe besteht aus den Items 21, 59 und 34. Die Items dieser Gruppe beinhalten Aussagen über die Existenz, Eindeutigkeit und Dauerhaftigkeit von Wahrheit. Die zweite Gruppe besteht aus den Items 33, 06 und 55. Die Items dieser Gruppe machen Aussagen über die Eindeutigkeit von Begriffen, die Glaubwürdigkeit von Aussagen in wissenschaftlichen Veröffentlichungen und die Konstanz von intellektuellen Leistungen über die Lebensspanne. Auf der Basis dieser Gruppierung wurde ein zweifaktorielles CFA-Modell mit den Items 21, 59 und 34 dem ersten Faktor zugehörend und die Items 33, 06 und 55 dem zweiten Faktor zugehörend erstellt. Dieses Modell erzielte einen guten Modellfit ($\chi^2 = 8.58$ ($df = 8$), $p = .38$, Bollen-Stine- $p = .55$, $\chi^2/df = .11$, CFI = .99, TLI = .99, SRMR = .0338, RMSEA = .018 (CI = [.000;.082], $p_{close} = .72$). Der χ^2 -Differenztest ergibt $\Delta\chi^2 = .04$ ($\Delta df = 1$), sodass hier nicht unbedingt von einer Verbesserung des Modells durch Hinzunahme eines weiteren Faktors ausgegangen werden kann. Ähnlich wie bei der ersten Skala gibt die Passung des zweifaktoriellen Modells auch hier einen Hinweis auf eine eventuell vorliegende Aufgliederung in zwei Facetten. Allerdings ist diese hier auch wiederum nicht mehr als eine begründete Vermutung, insbesondere, da der χ^2 -Differenztest nur ein geringfügige Verbesserung der Modellpassung durch das zweifaktorielle CFA-Modell andeutet. Allerdings ist im Gegensatz zur ersten Skala die Präzision von Cronbachs α deutlich in dem Bereich, in welchem von einer mehrdimensionalen Struktur ausgegangen werden kann.

Die dritte Skala wurde *Integratives Wissen* genannt. Die Benennung rührt von den Items 35, 18 und 54 her, die Aussagen über den Zweck und die Strukturierung von Texten sowie die Verbindung von neu Gelerntem und Vorwissen machen. Die übrigen Items machen Aussagen über die Gültigkeitsdauer aktueller wissenschaftlicher Befunde, die Schnelligkeit des Wissenserwerbs und die Eindeutigkeit der Aussagen von Dozenten. Das eindimensionale Messmodell führte zu einem guten Modellfit, ebenso das Messmodell mit den Restriktionen für τ -Äquivalenz. Allerdings konnte nicht von einem parallelen Messmodell ausgegangen werden.

Analog zu den vorhergehenden Faktoren ähneln sich wegen der τ -Äquivalenz auch die Werte für die Skalenreliabilität ρ und Cronbachs α mit jeweils .64. Auch wie bei den vorhergehenden Faktoren liegt die Item-Interkorrelation mit .24 in einem annehmbaren Bereich, aber die Präzision von Cronbachs α gibt mit .018 wieder einen deutlichen Hinweis auf eine mehrdimensionale Skala. Obwohl das eindimensionale Messmodell einen sehr guten Modellfit zeigt, liegen also dennoch Hinweise vor, dass sich auch diese Skala aus mehreren Facetten zusammensetzt.

Zur Prüfung dieser Vermutung wurde auch hier ein χ^2 -Differenztest berechnet. Die Items wurden zuerst auf der Basis ihres Inhaltes in Gruppen aufgeteilt. Die erste Gruppe bildeten die Items 35, 18 und 54, die zweite Gruppe die Items 61, 10 und 09. Es wurde ein zweifaktorielles CFA-Modell berechnet, welches einen guten Modellfit aufwies ($\chi^2 = 7.52$ ($df = 8$), $p = 0.48$, Bollen-Stine- $p = 0.59$, $\chi^2/df = .94$, CFI = 1.00, TLI = 1.00, SRMR = .0326, RMSEA = .000 (CI = [.000;.076], $p_{close} = .80$). Der χ^2 -Differenztest ergibt $\Delta\chi^2 = 9.95$ ($\Delta df = 1$). In diesem Fall liegt der $\Delta\chi^2$ -Wert über dem kritischen Wert von 3.84, sodass von einer Verbesserung des Modells durch die Hinzunahme eines weiteren Faktors ausgegangen werden kann. Bei der dritten Skala wird durch die Hinzunahme einer weiteren latenten Dimension eine deutliche Verbesserung der Modellpassung erzielt. Auch die Präzision von Cronbachs α ist hier deutlich höher als bei den ersten beiden Skalen, sodass sich diese Skala sich also in mehrere Facetten zu gliedern scheint.

Die vierte Skala wurde *Lernmethodik* genannt und bezieht sich auf den Umgang mit Lernmaterial. Mit Ausnahme des Items 25 beziehen sich alle Items auf Tätigkeiten, die den Umgang mit Gelerntem beziehen. Das Item 25 beinhaltet eine Aussage über Unterschiede in den kognitiven Fähigkeiten von Studenten und wurde deswegen aus der Skala ausgeschlossen. Das einfaktorielle CFA-Modell zeigte eine gute Passung, ebenso lag ein paralleles Messmodell vor. Aus diesem Grund ist auch die Skalenreliabilität ρ fast identisch mit Cronbachs α mit den Werten .59 bzw. .60. Allerdings liegt die Reliabilität der Skala zur Erfassung dieser Dimension unter dem akzeptablen Wert. Genau wie bei den vorhergegangenen Faktoren liegt die mittlere Item-Interkorrelation mit .27 im akzeptablen Bereich, aber die Präzision von Cronbachs α mit .20 gibt einen deutlichen Hinweis auf die Mehrdimensionalität der Skala. Allerdings war es bei diesem Faktor nicht möglich, die Items ihrem Inhalt entsprechend in Gruppen zu gliedern, die sich jeweils auf einen Aspekt beziehen. Jedes Item bezieht sich auf einen jeweils

eigenen Aspekt des Lernens, und ein entsprechendes Modell mit jeweils vier latenten Faktoren mit einem Indikator wäre nicht identifiziert (Kline, 1998).

Die letzte Skala wurde *Unsicherheit des Wissens* genannt. Die Benennung wurde gewählt, weil sie hauptsächlich Items enthält, die sich auf Unsicherheit beziehen (Items 27, 46, 42 und 30). Bei der Skalenbildung wurden das Item 14 und das Item 48 ausgeschlossen. Nach dem Ausschluss wies das eindimensionale Messmodell einen guten Modellfit auf und auch die Bedingungen für τ -Äquivalenz waren erfüllt, sodass von einem τ -äquivalenten Messmodell ausgegangen werden kann. Allerdings deutet bei dieser Skala sowohl die mittlere Interitemkorrelation mit einem Wert von .17 als auch die Präzision von Cronbachs α mit einem Wert von .028 auf eine mögliche Multidimensionalität der Skala hin. Entsprechend niedrig ist auch die Skalenreliabilität ρ bzw. Cronbachs α mit einem Wert jeweils .46. Wie es schon bei der Skala *Lernmethodik* war, ist es bei dieser Skala aufgrund der geringen Anzahl der Items nicht möglich, diese auf Basis ihres Inhalts den hypothetisch möglichen Facetten zuzuordnen, um damit ein CFA-Modell zu berechnen.

Es zeigt sich also bei allen fünf extrahierten Faktoren und den daraus abgeleiteten Skalen zur Messung dieser Faktoren ein kohärentes Muster. Die Skalen weisen bei einfaktoriellen konfirmatorischen Faktorenanalysen (ggf. nach geringfügigen Modifikationen durch Auslassen eines oder mehrerer Items) einen guten Modellfit und eine entsprechende Äquivalenz der Items auf. Allerdings muss im Hinblick auf die Ergebnisse einschränkend angemerkt werden, dass die letztendlich ermittelten Skalen an einer unabhängigen Stichprobe validiert werden sollten.

Die Reliabilität im Sinn der internen Konsistenz nach Cronbachs α der Skalen ist aber im geringen bis mittleren Bereich zu finden. Die geringen Reliabilitäten sind überdies in der Forschung zu epistemologischen Überzeugungen, die mithilfe des EBQ durchgeführt wird, gängig. Ein direkter Vergleich mit anderen vorliegenden Ergebnissen ist dagegen nur schwer möglich, und zwar wegen des angemerkten Fehlens von Analysen des EBQ auf Itemebene. Eine Vergleichsmöglichkeit bietet lediglich die Studie von Schraw et al. (2002), die ebenfalls Werte von Cronbachs α zwischen .74 und .61 für die ebenfalls von ihnen gefunden fünf Faktoren berichten. Auf die im Allgemeinen niedrigen Werte von Cronbachs α beim Gebrauch des EBQ haben aber auch Clarebout et al. (2001) hingewiesen. Schommer (1993a) begründet die niedrigen Reliabilitäten damit, dass die Versuchspersonen sich in einem Stadium des epis-

temologischen Übergangs befinden. In diesem Übergang erkennen die Versuchspersonen, dass es andere epistemologische Standpunkte gibt und sind deswegen bei der Wahl der Antwort auf ein Item inkonsistent, was zu niedrigen Reliabilitäten führt. Die dauerhaft niedrigen Reliabilitätskennwerte würden aber in dieser Auslegung bedeuten, dass dieses Stadium des epistemologischen Übergangs ein andauernder Zustand wäre. Überdies wurde die Untersuchung von Schommer (1993a) an Schülern durchgeführt, während die vorliegende Untersuchung auf Studierende und Berufstätige abzielte. Eine Interpretation der niedrigen Reliabilitäten als Phase des Übergangs zwischen verschiedenen epistemologischen Sichtweisen würde bedeuten, dass dieser Übergang nicht nur bei Schülern, sondern auch bei Studierenden und Berufstätigen zu finden ist und damit sehr lange andauert, was aber der Vorstellung von epistemologischen Überzeugungen als Traits widerspricht, da Traits über einen angemessenen Zeitrahmen stabil sein sollten (Asendorpf, 1999).

Eine alternative Erklärung als eine Phase des epistemologischen Übergangs gestatten die niedrigen Reliabilitätskennwerte zusammen mit den am unteren Ende des akzeptablen Bereichs liegenden mittleren Interitemkorrelationen und den hohen Werten der Präzision von Cronbachs α . Demnach handelt es sich bei diesen Faktoren und den hieraus abgeleiteten Skalen nicht um homogene Konstrukte. Vielmehr scheint es sich um Skalen zu handeln, die sich in Facetten gliedern lassen. Dies ist den Befunden aus dem Persönlichkeitsbereich der Big Five ähnlich, bei denen die fünf Faktoren breitere Konstrukte darstellen, die sich wiederum in Facetten unterteilen lassen (DeRaad, 2000). Die Big Five gliedern sich aber nicht nur in Facetten, diese Persönlichkeitsdimensionen lassen sich zu den beiden höheren Faktoren Alpha und Beta zusammenfassen (Digmann, 1997). Damit liegt ein hierarchisches Persönlichkeitsmodell vor. Solche Modelle stellen eine allgemeine Klasse von Persönlichkeitsmodellen dar (Schweizer, Moosbrugger & Schermelleh-Engel, 2003), bei der mehrere latente Dimensionen hierarchisch angeordnet sind. Man spricht je nach Gestaltung des Modells von einem *Higher-order*-Faktormodell bzw. einem Modell hierarchischer Faktoren. Dass solche Modelle bei epistemologischen Überzeugungen Anwendung finden können, wurde von Wood und Kardash (2002) sowie Bromme (2005) angemerkt. Ein hierarchisches Modell⁴¹ würde somit be-

⁴¹ Der Unterschied zwischen einem *Higher-order*-Faktormodell und einem hierarchischen Faktor-Modell ist, dass bei einem *Higher-order*-Faktormodell die manifesten Variablen nur von den latenten Faktoren auf der ersten Ebene des Modells beeinflusst werden, die latenten Faktoren auf der ersten Ebene werden nur durch die latenten Faktoren auf der zweiten Ebene beeinflusst usw. Die Faktoren auf der zweiten Ebene können bei einem *Higher-order*-Faktormodell also nur indirekt vermittelt durch die latenten Faktoren der ersten Ebene auf die manifesten Variablen wirken. Bei einem hierarchischen Faktormodell hingegen wirken sowohl die Faktoren der zweiten Ebene als auch die Faktoren der ersten Ebene direkt auf die manifesten Variablen (Yung et al., 1999). Diese Unterscheidung ist in erster Linie methodisch zu sehen und hängt nur indirekt mit einer Hierarchisierung

deuten, dass die gefundenen Faktoren jeweils Faktoren einer höheren Ebene darstellen und dass sich diese – analog den zuvor erwähnten Facetten der Big Five – in untergeordnete Faktoren untergliedern lassen.

Die mögliche Erklärung der niedrigen Reliabilitätskennwerte durch eine Facettierung lässt sich auch unter dem Aspekt der Homogenität der Items betrachten. Homogenität bezieht sich auf die inhaltliche Einheitlichkeit der Items⁴² (Amelang & Zielinski, 2002). Homogene Items beschreiben einen Aspekt einer Merkmalsdimension. Es wurde bereits diskutiert, dass die Items inhaltlich teilweise keinen gemeinsamen Bezug haben, was sich auch in dem niedrigen KMO-Maß bemerkbar macht. Ein Beispiel für eine Skala mit niedriger Reliabilität und inhomogenen Items bietet der Faktor *Unsicherheit des Wissens*. Dieser besteht aus den Items „Die erfolgreichsten Leute haben entdeckt, wie sie ihre Lernfähigkeit verbessern können“, „Es macht mir Spaß, über Dinge nachzudenken, über die sich Fachleute nicht einigen können“, „Das einzig Gewisse ist die Ungewissheit“, „Ich mag Filme mit offenem Ende nicht“ und „Auch der Rat von Fachleuten ist oft anzuzweifeln“. Teilweise ähneln sich die Items nicht sehr. Bei der Benennung des Faktors wurde schon angesprochen, dass sich die Items auf Aspekte der Unsicherheit, der Präferenz von Denkprozessen über schwierige Aufgaben und der Erkennbarkeit der Lernfähigkeit beziehen. Obwohl sich die Mehrheit auf Aspekte der Unsicherheit beziehen, ist nicht erkennbar, dass die Items *einen* Aspekt *einer* Merkmalsdimension beschreiben. Vielmehr werden unterschiedliche Merkmalsdimensionen beschrieben bzw. bei der Merkmalsdimension Unsicherheit werden unterschiedliche Aspekte derselben beschrieben. Letztendlich lässt sich somit trotz des passenden Messmodells keine definitive Aussage über die Homogenität der Items tätigen, die empirischen Befunde liefern letztendlich nur Indizien. Trotz der inferenzstatistischen Absicherung des Messmodells für den Faktor *Unsicherheit des Wissens* sprechen die mangelnde Reliabilität, die Höhe der Präzision von Cronbachs α und die Referenz der Items unterschiedlicher Merkmalsdimensionen für die Annahme, dass es sich um einen Faktor handelt, dessen Items keine faktorielle Validität haben.

Allerdings kann auch nicht sicher auf eine mögliche hierarchische Struktur geschlossen werden. So wurde bei den informellen Berechnungen bei der Diskussion der Ergebnisse nur beim

von Eigenschaften wie in den bekannten hierarchischen Persönlichkeitsmodellen zusammen. Da hierarchische Faktormodelle methodologisch einige Vorteile bieten (Gignac, 2007) und da im Rahmen dieser Arbeit hauptsächlich auf die Idee, welche den hierarchischen Modellen zugrunde liegt, Bezug genommen wird, wird im Folgenden nur von hierarchischen Modellen gesprochen.

⁴² Genauer betrachtet bezieht sich Homogenität auf die formale und inhaltliche Einheitlichkeit der Items (Amelang & Zielinski, 2002). Da das Itemformat als formaler Gestaltungspunkt für alle Items gleich ist, kann die Betrachtung der formalen Einheitlichkeit im Folgenden unberücksichtigt bleiben.

dritten Faktor ein deutlicher Hinweis auf eine eventuell vorhandene mehrdimensionale Struktur gefunden. Bei den ersten beiden Faktoren war die Steigerung der Modellpassung, ausgedrückt durch den χ^2 -Differenztest, nicht ausreichend. Bei den letzten beiden Faktoren konnte inhaltlich keine mehrfaktorielle Struktur zum Testen der Modellpassung abgeleitet werden. Zudem stellen sich methodische Probleme, darunter fallen insbesondere äquivalente Modelle. Zwischen hierarchischen Modellen und Modellen mit korrelierten Faktoren kann statistisch nicht unterschieden werden, da diese durch die gleiche Kovarianzmatrix beschrieben werden (Bühner, 2006). Z. B. ist ein Modell mit zwei korrelierenden Faktoren äquivalent zu einem Modell mit einem Second-order-Faktor und zwei First-order-Faktoren. Auf der Grundlage der Fitindizes können die Modelle nicht unterschieden werden. Hierzu bedarf es weiterer Methoden, wie z. B. der Schmid-Leimann-Transformation, die es gestatten, den Einfluss der Second-order- und First-order-Faktoren jeweils getrennt voneinander auf die manifesten Variablen und deren psychologische Bedeutsamkeit beurteilen zu können (Schmid & Leimann, 1957, Wolf & Preising, 2005, Gignac, 2007). Dazu bedarf es wiederum theoretischer Modelle, welche Items welchen Faktoren zugeordnet werden können. Diese Modelle können induktiv mittels explorativer Faktorenanalyse entwickelt werden, wobei es anschließend zur Überprüfung der entwickelten Modelle einer neuen Stichprobe von Versuchspersonen bedarf. Die durchgeführte EFA des EBQ hat aber ergeben, dass die Entwicklung von Faktormodellen mit dessen Items auf einige Schwierigkeiten stößt.

An dieser Stelle kann eine abschließende Bewertung der zweiten und dritten Forschungsfrage erfolgen. Diese Forschungsfragen betreffen die faktorielle Struktur des EBQ bei einer Analyse auf Itemebene sowie die Frage, ob es möglich ist, Skalen zur reliablen Messung der Faktoren abzuleiten. Die Items des Schommer-Fragebogens bilden eine 5-faktorielle Struktur, wobei diese aber noch einer weiteren Validierung bedarf. Insbesondere muss geklärt werden, ob die Faktoren sich weiter in Facetten aufgliedern oder ob es sich bei epistemologischen Überzeugungen um Konstrukte handelt, die sich mit der Zeit verändern, wie es von Schommer (1993a) vorgeschlagen wurde. Reliable Skalen zur Erfassung epistemologischer Überzeugungen konnten nicht abgeleitet werden.

Die große Anzahl an Items, die vorab von der Faktorenanalyse ausgeschlossen werden mussten, sowie die mangelnde Reliabilität der abgeleiteten Skalen machen es auch erforderlich, die Eignung der Items zur Erfassung epistemologischer Überzeugungen zu klären. Auf keinen Fall sollte der EBQ in der vorliegenden Form weder nach den Vorgaben von Schommer noch

nach den hier abgeleiteten Skalen zur Erhebung epistemologischer Überzeugungen eingesetzt werden.

5. Ergebnisse zum EPI

5.1 Faktorenanalytische Auswertung des EPI

Das Ziel dieses Kapitels ist die Beantwortung der ersten vier der insgesamt fünf Forschungsfragen zum EPI-Fragebogen. Diese beschäftigen sich hauptsächlich mit der Faktorenstruktur des EPI-Fragebogens. Fragen der Konstruktvalidität, die sich auf die fünfte Forschungsfrage beziehen, werden im darauffolgenden Kapitel besprochen. Behandelt werden hier die folgenden Forschungsfragen:

- *Lässt sich die Faktorenstruktur des EPI-Fragebogens mithilfe einer konfirmatorischen Faktorenanalyse bestätigen?*
- *Lassen sich die einzelnen Skalen des EPI-Fragebogens mithilfe einer konfirmatorischen Faktorenanalyse bestätigen und welche Ähnlichkeiten und Unterschiede gibt es zwischen diesen Ergebnissen und dem FEE von Moschner und Gruber (im Druck)?*
- *Lassen sich ggf. notwendige Modifikationen der Skalen des EPI mithilfe einer Kreuzvalidierung absichern?*
- *Welche psychometrischen Eigenschaften besitzen die Skalen des EPI-Fragebogens?*

5.1.1 Vorbereitung der Daten

Der Fragebogen wurde insgesamt zwei verschiedenen Stichproben vorgelegt. Die erste Stichprobe stellt die eigentliche Analytestichprobe dar, die zweite die Validierungsstichprobe. In der ersten Stichprobe wurde der Fragebogen von 419 Studenten bearbeitet. Die Probanden waren Lehramtsstudierende der Universität des Saarlandes. Von den ursprünglichen 419 Fällen wurden 128 Fälle ausgeschlossen, bei denen das Alter bzw. das Geschlecht nicht angegeben und bei denen mehr als fünf fehlende Werte bei den EPI-Items vorhanden waren. Die Auswahl des Kriteriums von fünf fehlenden Werten beruht auf der Annahme, dass mehr als 10 % fehlende Antworten auf ein willkürliches Antwortverhalten hindeuten können und damit nicht für eine Auswertung geeignet sind, wobei dem Probanden gleichzeitig aber ein gewisser Anteil an ausgelassenen Antworten (z. B. durch Übersehen eines Items) zugestanden wird. Bei den verbleibenden 291 Fällen wurden die fehlenden Werte mittels des EM-Algorithmus geschätzt (Allison, 2012, 2009, Lüdtke et al., 2007). Die Auswertung erfolgte mithilfe der beiden Programme SPSS 15 und Amos 16. Die Altersverteilung der Probanden aus der Gesamtstichprobe ist in Tabelle 60 zu finden.

Tabelle 60: Deskriptive Statistik der Altersverteilung in Jahren - Analysestichprobe

Geschlecht	<i>n</i>	Range	Min	Max	<i>M</i>	<i>SD</i>
männlich	119	22	19	41	22.83	4.19
weiblich	172	29	18	47	21.15	4.03
gesamt	291	29	18	47	21.84	4.17

Der überwiegende Anteil der Versuchsteilnehmer war weiblich, was sich aus der Rekrutierung aus Lehramtsstudiengängen erklären lässt, da es hier einen größeren Frauenanteil gibt. In Tabelle 61 ist die Verteilung der Probanden auf die Fächergruppen zu finden. Die Gruppierung erfolgte nach der Wahl des gewählten ersten Faches für das Lehramtsstudium. Die Fächer wurden in Gruppen zusammengefasst, bis auf die Lehrämter für Wirtschaftspädagogik sowie die Fächer Deutsch und Sport, die nicht eingruppiert wurden. In der Gruppe Fremdsprachen sind alle wählbaren Fremdsprachen erfasst, in der Gruppe Naturwissenschaften die Fächer Biologie, Physik, Chemie und Informatik, in die Gruppe Gesellschafts-/Kulturwissenschaften fallen die Fächer Geografie, katholische und evangelische Theologie, Philosophie, Geschichte sowie Wirtschaftswissenschaft.⁴³

Tabelle 61: Verteilung der Probanden auf Fächergruppen - Analysestichprobe

Geschlecht	gesamt		männlich		weiblich	
	Häufigkeit	%	Häufigkeit	%	Häufigkeit	%
Fach						
Wirtschaftspädagogik	17	5.8	11	9.2	6	3.5
Deutsch	61	21.0	18	15.1	43	25.0
Fremdsprachen	79	27.1	21	17.6	58	33.7
Sport	19	6.5	11	9.2	8	4.7
Naturwissenschaften	68	23.4	41	34.5	27	15.7
Gesellschafts- /Kulturwissenschaften	33	11.3	13	10.9	20	11.6
Musik/Kunst	8	2.7	1	.8	7	4.1
Keine Angabe	6	2.1	3	2.5	3	1.7
Gesamt	291	100	119	100	172	100

Um die Daten auf ihre Eignung zur Durchführung einer CFA zu untersuchen, wurden zuerst die Schiefe und die Wölbung aller EPI-Items untersucht, die Werte sind in sind in Tabelle 61 dargestellt.⁴⁴

⁴³ Der Lehramts-Studiengang Wirtschaftspädagogik war zu dem Zeitpunkt der Erhebung vom Fach Wirtschaftswissenschaft getrennt und führte zu dem Abschluss Diplom-Wirtschaftspädagoge.

⁴⁴ An dieser Stelle sind alle Items mit ihren Schiefen und Wölbungen aufgeführt. Da für die Analysen der einzelnen Skalen jeweils eine Teilmenge der in Tabelle 62 aufgeführten Items benutzt wird, wird dort auf eine tabellarische Darstellung der Schiefe und Wölbung verzichtet und nur jeweils der Wert des Mardia-Tests angegeben. Die Verteilungsparameter der Items können bei Bedarf in dieser Tabelle nachgesehen werden.

Tabelle 62: Verteilungsform der EPI-Items – Analysestichprobe

Variable	Schiefe	CR	Wölbung	CR
Item 01	.68	4.74	-.27	-.92
Item 02	.71	4.96	.19	.66
Item 03	1.51	1.53	2.53	8.80
Item 04	.63	4.35	-.46	-1.60
Item 05	.05	.35	-1.04	-3.61
Item 06	-.73	-5.09	-.18	-.61
Item 07	-.59	-4.07	.53	1.83
Item 08	-.07	-.51	-1.17	-4.09
Item 09	.46	3.22	-.34	-1.20
Item 10	.67	4.69	-.22	-.78
Item 11	-1.03	-7.14	1.20	4.17
Item 12	-.95	-6.64	.73	2.55
Item 13	-.86	-5.96	-.22	-.76
Item 14	-.33	-2.28	-.95	-3.32
Item 15	.71	4.96	-.12	-.43
Item 16	-.51	-3.52	.01	.02
Item 17	-.16	-1.13	-.72	-2.50
Item 18	-.02	-.14	-.94	-3.28
Item 19	.10	.68	-.96	-3.35
Item 20	.73	5.11	-.04	-.15
Item 21	-.44	-3.06	-.05	-.18
Item 22	-.57	-3.96	-1.04	-3.62
Item 23	-.80	-5.55	.07	.23
Item 24	.38	2.63	-.92	-3.20
Item 25	-.83	-5.76	.06	.21
Item 26	.86	6.01	.64	2.23
Item 27	-.60	-4.16	.54	1.89
Item 28	-.62	-4.31	.78	2.71
Item 29	1.05	7.32	.44	1.52
Item 30	-.97	-6.76	.89	3.09
Item 31	-.91	-6.36	.23	.82
Item 32	-.65	-4.52	-.22	-.76
Item 33	-1.20	-8.37	1.95	6.80
Item 34	-.49	-3.43	-.07	-.24
Item 35	-.93	-6.48	.92	3.20
Item 36	-.81	-5.63	.64	2.24
Item 37	-.02	-.16	-.57	-1.99
Item 38	.33	2.32	-.86	-2.98
Item 39	-.97	-6.73	1.01	3.52
Item 40	-.60	-4.18	-.67	-2.34
Item 41	-.08	-.55	-1.14	-3.96
Item 42	.46	3.23	-.53	-1.84
Item 43	-.58	-4.03	-.63	-2.18
Item 44	-.91	-6.35	-.27	-.93
Item 45	-.54	-3.75	.21	.74

Ergebnisse zum EPI

Variable	Schiefe	CR	Wölbung	CR
Item 46	-.57	-4.00	-.18	-.63
Item 47	-1.65	-11.51	2.92	10.15
Item 48	-.99	-6.88	1.46	5.07
Item 49	-.06	-.45	-.58	-2.01
Item 50	-.60	-4.14	-.03	-.12
Item 51	-.72	-5.04	.96	3.35
Item 52	1.82	12.65	3.98	13.87
Item 53	-.23	-1.60	-.79	-2.75
Mardia-Test			337.23	37.67

Keines der Items zeigt einen Wert über den Grenzen von West et al. (1995). Allerdings weisen einige Items doch eine deutliche Schiefe bzw. Wölbung auf, wie z. B. das Item 47 und 52. Einige *critical ratios* (CR) weisen auf eine signifikante Abweichung der Verteilungsform von der Normalverteilung hin. Der Mardia-Test, der auf eine multivariate Normalverteilung prüft, ist signifikant ($M = 337.23$, $CR = 37.67$).

Die zweite Stichprobe bestand zu einen aus 154 Versuchspersonen, die aus Studierenden der Erziehungswissenschaft der Universität des Saarlandes rekrutiert wurden, und zum anderen aus Versuchspersonen, die den Fragebogen im Internet ausgefüllt haben. Die Erhebung der Daten fand während des Sommersemesters 2009 statt, die Onlineversion des Fragebogens wurde in der Zeit von Februar bis August 2009 auf der Internetseite <http://www.internetpsychologie.net/> dargeboten. Von den ursprünglichen 154 Fällen wurden 28 Fälle ausgeschlossen, bei denen das Alter bzw. das Geschlecht nicht angegeben wurde und bei denen mehr als fünf fehlende Werte bei den EPI-Items vorhanden waren. So verbleiben 126 Fälle, bei denen eventuell noch vorhandene fehlende Werte durch den EM-Algorithmus geschätzt wurden. Die Altersverteilung der Validierungsstichprobe ist in Tabelle 63 zu finden.

Tabelle 63: Deskriptive Statistik der Altersverteilung in Jahren – Validierungsstichprobe

Geschlecht	<i>n</i>	Range	Min	Max	<i>M</i>	<i>SD</i>
männlich	48	35	19	54	28.81	10.35
weiblich	78	25	19	44	23.38	5.29
gesamt	126	35	19	54	25.45	8.03

Die Verteilung der Validierungsstichprobe auf die unterschiedlichen Fächergruppen ist in Tabelle 64 dargestellt. Auch bei der Validierungsstichprobe finden sich mehr Frauen als Männer. Überdies war unter den Teilnehmern des Online-Fragebogens eine Person, die eine

berufliche Tätigkeit statt eines Studiums angab. In Tabelle 65 sind die Verteilungsformen der Items aus der Analysestichprobe zu finden.

Tabelle 64: Verteilung der Probanden auf Fächergruppen – Validierungsstichprobe

Geschlecht Fach	Gesamt		Männlich		Weiblich	
	Häufigkeit	%	Häufigkeit	%	Häufigkeit	%
Beruf	1	0.8	1	2.1	0	0
Deutsch	16	12.7	2	4.2	14	17.9
Fremdsprache	22	17.5	7	14.6	15	19.2
Gesellschafts- /Kulturwissenschaften	71	56.3	29	60.4	42	53.8
Naturwissenschaften	8	6.3	5	10.4	3	3.8
Sport	2	1.6	0	0	2	2.6
Keine Angabe	6	4.8	4	8.3	2	2.6
Gesamt	126	100.0	48	100.0	78	100.0

Tabelle 65: Verteilungsform der EPI-Items – Validierungsstichprobe

Variable	Schiefe	CR	Wölbung	CR
Item 01	-.72	-3.29	-.88	-2.02
Item 02	.28	1.27	-1.00	-2.30
Item 03	.84	3.84	.29	.66
Item 04	.63	2.87	-.60	-1.37
Item 05	-.83	-3.80	-.28	-.65
Item 06	-.63	-2.88	-.33	-.75
Item 07	-2.21	-1.13	7.43	17.02
Item 08	.21	.95	-1.27	-2.90
Item 09	-.07	-.31	-.84	-1.92
Item 10	.90	4.14	.50	1.15
Item 11	-1.29	-5.93	2.80	6.42
Item 12	-.67	-3.06	-.20	-.46
Item 13	-.77	-3.52	-.39	-.89
Item 14	-.78	-3.58	-.60	-1.38
Item 15	.56	2.56	-.63	-1.44
Item 16	-.81	-3.71	1.14	2.60
Item 17	-.18	-.81	-1.00	-2.29
Item 18	-.34	-1.57	-.97	-2.22
Item 19	-.45	-2.08	-.85	-1.95
Item 20	.63	2.89	-.08	-.19
Item 21	-.83	-3.81	.07	.16
Item 22	-.23	-1.04	-1.10	-2.51
Item 23	-.65	-2.99	.29	.67
Item 24	.78	3.58	-.07	-.15
Item 25	-.77	-3.53	.06	.14
Item 26	.74	3.41	.50	1.15
Item 27	-.54	-2.49	.36	.82
Item 28	-.49	-2.22	-.01	-.02

Ergebnisse zum EPI

Variable	Schiefe	CR	Wölbung	CR
Item 29	.45	2.05	-1.06	-2.43
Item 30	-1.11	-5.10	1.92	4.39
Item 31	-.70	-3.19	.24	.54
Item 32	-.90	-4.13	1.13	2.58
Item 33	-.11	-.49	-.73	-1.66
Item 34	-.98	-4.49	1.16	2.66
Item 35	-1.24	-5.70	2.86	6.55
Item 36	-.77	-3.54	.47	1.08
Item 37	-.79	-3.60	.25	.56
Item 38	.40	1.83	-.83	-1.91
Item 39	-.90	-4.14	1.80	4.13
Item 40	-1.11	-5.06	.78	1.80
Item 41	.03	.13	-.87	-1.98
Item 42	.02	.07	-1.15	-2.63
Item 43	-.54	-2.47	-.61	-1.41
Item 44	-.42	-1.94	-.80	-1.84
Item 45	-.75	-3.42	.65	1.48
Item 46	-.72	-3.29	.32	.74
Item 47	-1.39	-6.35	2.18	5.00
Item 48	-.93	-4.28	1.91	4.38
Item 49	-.34	-1.54	-.53	-1.22
Item 50	-.90	-4.12	1.37	3.15
Item 51	-1.24	-5.69	3.16	7.24
Item 52	1.64	7.50	2.81	6.44
Item 53	-.24	-1.10	-.70	-1.61
Mardia-Test			127.24	9.35

In der Validierungsstichprobe zeigte das Item 07 einen Wert über den Grenzen von West et al. (1995). Der Mardia-Test auf multivariate Normalverteilung ist auch in der Validierungsstichprobe signifikant.

An dieser Stelle wird an das in den Methoden beschriebene Vorgehen erinnert, dass bei den Modellen standardmäßig immer die mittels des Bollen-Stine-Bootstraps geschätzte Wahrscheinlichkeit für den gegebenen χ^2 -Wert beachtet wird. Liegen Abweichungen der Daten von der Normalverteilung bzw. der multivariaten Normalverteilung vor, so liefert dieser Wert ein genaueres Bild über die Wahrscheinlichkeit des χ^2 -Werts. Liegen keine Abweichung vor, sind die direkt berechnete Wahrscheinlichkeit und die Wahrscheinlichkeit aus dem Bollen-Stine-Bootstrap nahezu identisch.

5.1.2 CFA des Gesamtmodells mit sieben Faktoren

Die erste Forschungsfrage beschäftigt sich mit der von Moschner und Gruber vorgeschlagenen Faktorenstruktur des EPI. Insbesondere steht dabei die Frage im Raum, ob die vorgeschlagene Faktorenstruktur sich mithilfe einer konfirmatorischen Faktorenanalyse bestätigen lässt. Dazu wird im ersten Schritt das Gesamtmodell mit den sieben Faktoren des EPI-Fragebogens geprüft. Die Skalierer waren dabei jeweils das erste Item einer Skala, dessen Faktorladung auf eins fixiert wurde. Weiterhin wurden alle möglichen Kovarianzen zwischen den Faktoren geschätzt. Die Berechnung der Fitindizes ergab folgende Werte: $\chi^2 = 2654.85$ ($df = 1304$), $p < .01$, Bollen-Stine- $p = .001$, $\chi^2/df = 2.04$, CFI = .65, TLI = .63, SRMR = .0839 und RMSEA = .060, CI = [.057;.063], $p_{close} < 001$.

Der Bollen-Stine- p Wert, der CFI-Wert und TLI- Wert deuten auf eine inakzeptable Modellpassung hin, die restlichen Indizes bewegen sich im Rahmen akzeptabler Modellpassung. Das Gesamtmodell muss also in der gegenwärtigen Spezifikation verworfen werden. Gemäß der in Kapitel 3.3 dargestellten Vorgehensweise wird jeder Faktor separat in einem CFA-Modell betrachtet, um, sofern dies möglich ist, durch entsprechende Modellrevision ein passendes Messmodell zu erhalten.

5.1.3 Prüfung der einzelnen Faktoren

5.1.3.1 Faktor Reflexive Natur des Wissens

Das erste Modell RN (siehe Tabelle 68, die Tabellen mit den Fitindizes werden der Übersichtlichkeit wegen für alle Faktoren des EPI-Fragebogens am Schluss eines Kapitels aufgeführt) erzielte eine mäßige Anpassung an die Daten.⁴⁵ Aufgrund inhaltlicher Überlegungen wurden Items 27 ausgeschlossen, obwohl dies durch die Residuen nicht nahegelegt wurde, lediglich der Modifikationsindex gab Hinweise auf eine Korrelation der Fehlervariablen von Item 27 und Item 28. Der Wert des Modifikationsindex lag über einem Wert von 5, die geschätzte Korrelation der Fehlervariablen würde allerdings nur .08 betragen. Inhaltlich betrachtet bezie-

⁴⁵ In den Ergebnistabellen der konfirmatorischen Faktorenanalysen wird jeweils das Modell mittels seiner Abkürzung aufgeführt. Das ursprünglich getestete Modell wird jeweils nur mit seiner Abkürzung aufgeführt, die Modelle, die modifiziert wurden, erhalten hinter der Abkürzung eine arabische Ziffer, welche den Modifikations-schritt angibt. Das finale Modell erhält die Kennzeichnung „final“ nach dem Modellkürzel. Die Fitindizes der Validierungsstichproben werden mit den Kennzeichnung S1 bis S10 nach dem Modellkürzel gekennzeichnet. Die Fitindizes für die Testung der kongenerisch, τ -äquivalenten und parallelen Modellen sind mit entsprechenden Abkürzungen versehen, die Fitindizes für das an der Validierungsstichprobe getestete Modell ist mit dem Kürzel „Valid.“ gekennzeichnet. Die Fitindizes für das kongenerische Modell sind gleich denen des finalen Modells – da die Modelle gleich sind, diese sind nur der Übersichtlichkeit halber erneut aufgeführt.

hen sich die beiden Items auf die Veränderung von Wissensstrukturen, wenn das bisher vorhandene Wissen mit neuem Wissen angereichert wird. Die inhaltliche Ähnlichkeit von Item 27 und 28 kann zu einem korrelierenden Messfehler führen, welcher äquivalent zu einem Methodenfaktor sind. Diese stellen eine weitere Varianzquelle dar, die neben dem zugrunde liegenden Konstrukt die manifesten Variablen bedingen, es handelt sich somit um eine weitere systematische Varianzquelle. Die Verbesserung des Modellfits stützt damit die Vermutung, dass die beiden Items neben dem eigentlichen Faktor eine weitere latente Varianzquelle gemeinsam haben. Die Modellanpassung verbesserte sich durch den Ausschluss von Item 27. Das Modell RN final erzielte eine sehr gute Modellanpassung.

Die Kreuzvalidierung erweist sich dagegen als problematisch. Das Modell wies in der Validierungsstichprobe eine sehr schlechte Modellpassung auf, lediglich der SRMR und der p -Wert des Bollen-Stine-Bootstrap zeigen Werte, die auf eine akzeptable Modellpassung schließen lassen. In den zehn Teilstichproben der Analysestichprobe hingegen wies das Modell bis auf wenige Ausnahmen Fitwerte auf, welche auf eine gute Modellpassung schließen lassen. Es ist somit fraglich, ob der Ausschluss von Item 27 gerechtfertigt war. Um dies zu überprüfen, wurde statt des Items 27 das Item 28 aus dem Faktormodell genommen. Dieses geänderte Messmodell wurde in der Analysestichprobe neu berechnet. Das Modell wies eine weitgehend akzeptable Modellpassung auf ($\chi^2 = 5.77$ ($df = 2$), $p = .06$, Bollen-Stine- $p = .09$, $\chi^2/df = 2.89$, CFI = .98, TLI = .94, SRMR = .0308, RMSEA = .081, CI = [.000;.161], $p_{close} = .19$). Das Messmodell ohne Item 28 kann also zu den Daten passend angenommen werden. In der Validierungsstichprobe erzielte das Messmodell ohne Item 28 ebenso eine gute Modellpassung ($\chi^2 = .15$ ($df = 2$), $p = .93$, Bollen-Stine- $p = .94$, $\chi^2/df = .07$, CFI = 1.00, TLI = 1.16, SRMR = .0083, RMSEA = .000, CI = [.000;.053], $p_{close} = .95$). Das Modell kann demnach nicht ohne Weiteres auf andere als die hier vorliegende Stichprobe übertragen werden. Der vorgenommene Ausschluss von Item 28 legt nahe, dieses Item statt Item 27 zu entfernen, was zuvor aber einer Prüfung an einer weiteren unabhängigen Stichprobe bedarf.

Der Mardia-Test für die vier Items 07, 28, 58 und 51 des finalen Messmodells ist signifikant ($M = 14.48$, $CR = 17.82$), keines der Items zeigt eine Schiefe oder eine Wölbung über den Grenzwerten von West et al. (1995). Bei der Interpretation der Parameterwerte ist daher mit Unterschätzungen der Standardfehler der Parameter und mit überhöhten Signifikanzen zu

rechnen. Die standardisierten Faktorladungen weisen alle eine substantielle Höhe auf und sind auch auf dem .001%-Niveau signifikant.⁴⁶

Tabelle 66: Parameterwerte für Modell RN 1

Item	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
Item 07	1.00	.67			
Item 28	1.00	.61	.13	7.83	***
Item 48	.80	.52	.12	6.95	***
Item 51	1.12	.74	.14	8.33	***

Anmerkung: ***: $p < .001$

Das Modell RN final zeigt sowohl unter den Restriktionen der τ -Äquivalenz als auch der Parallelität bei fast allen Fitindizes eine gute Modellpassung.⁴⁷ Lediglich der RMSEA des parallelen Modells fällt in den Bereich einer akzeptablen Modellpassung. Auffällig beim parallelen Modell ist ebenfalls, dass der nicht durch die Bollen-Stine-Prozedur korrigierte p-Wert stark von dem korrigierten abweicht. Dies kann durch die Abweichung der Daten von der multivariaten Normalverteilung bedingt sein. Das Modell RN final stellt somit ein paralleles Messmodell für die latente Variable *Reflexive Natur des Wissens* dar. Cronbachs α ist somit ein sinnvoll einzusetzender Schätzwert für die Reliabilität der Skala. Wie Tabelle 67 zeigt, weichen Cronbachs α und Skalenreliabilität ρ für die entsprechend restringierten Messmodelle auch nur minimal voneinander ab.

Tabelle 67: Reliabilitätsschätzungen der Skala Reflexive Natur des Wissens

Reliabilitäten
Skalenreliabilität ρ kongenerisch = .73
Skalenreliabilität ρ tau = .72
Skalenreliabilität ρ parallel = .72
Cronbachs α = .73

Die Reliabilität der Skala liegt in dem Bereich, der typisch ist für Skalen zur Messung epistemologischer Überzeugungen, wobei besonders die Äquivalenz der Messungen zu betonen ist. Wegen der Parallelität ist jedes Item gleichermaßen gut zur Messung der latenten Variablen geeignet. Ob sich das parallele Messmodell auf andere Stichproben übertragen lässt, bedarf

⁴⁶ Die standardisierte Faktorladung für die Skaliervariable kann zwar berechnet, aber nicht auf Signifikanz geprüft werden.

⁴⁷ An dieser Stelle sei kurz darauf hingewiesen, dass die Angabe des kongenerischen Modells in Tabelle 68 redundant ist, da es Modell dem Modell RN final entspricht. Das kongenerische Modell wird in jeder Tabelle in der drittletzten Spalte immer aus Gründen der Übersichtlichkeit dargestellt.

weiterer Untersuchungen. Wegen der Parallelität⁴⁸ sind die Skalenreliabilität und Cronbachs α numerisch annähernd gleich. Die Höhe der Reliabilität liegt mit rund .73 im niedrigen, aber noch akzeptablen Bereich. Zusammenfassend kann festgehalten werden, dass mit den Items 07, 28, 48 und 51 eine reliable Erfassung der latenten Variable *Reflexive Natur des Wissens* gegeben ist. Allerdings ist wegen der fehlenden Modellpassung des Messmodells in der Validierungsstichprobe keine überzeugende Skalenkonstruktion zur Erfassung dieser Dimension epistemologischer Überzeugungen gelungen.

Tabelle 68: Ergebnisse der Skala Reflexive Natur des Wissens

Modell	N	χ^2 (df)	p	Bollen-Stine-p	χ^2/df	CFI	TLI	RMSEA	CI (90 %)	p_{close}	SRMR
RN	291	12.76 (5)	.03	.08	2.55	.97	.95	.073	.023;.124	.18	.0342
RN final	291	.39 (2)	.82	.85	.20	1.00	1.02	.000	.000;.069	.91	.0074
RN S1	207	.86	.65	.69	.43	1.00	1.02	.000	.000;.108	.77	.0123
RN S2	221	.34	.85	.86	.17	1.00	1.03	.000	.000;.075	.90	.0078
RN S3	217	.58	.75	.80	.29	1.00	1.03	.000	.000;.093	.84	.0112
RN S4	212	.04	.98	.98	.02	1.00	1.03	.000	.000;.000	.99	.0027
RN S5	221	1.22	.54	.63	.61	1.00	1.01	.000	.000;.116	.69	.0153
RN S6	215	.31	.86	.88	.16	1.00	1.03	.000	.000;.073	.91	.0078
RN S7	217	1.87	.39	.47	.93	1.00	1.00	.000	.000;.132	.56	.0197
RN S8	221	.32	.85	.88	.16	1.00	1.03	.000	.000;.073	.91	.0073
RN S9	219	1.29	.53	.56	.64	1.00	1.01	.000	.000;.118	.68	.0150
RN S10	214	.94	.63	.69	.47	1.00	1.02	.000	.000;.109	.76	.0143
RN kong.	291	.39 (2)	.82	.82	.20	1.00	1.02	.000	.000;.069	.91	.0074
RN tau	291	6.73 (5)	.24	.50	1.35	.99	.99	.035	.000;.094	.51	.0337
RN paral- lel	291	14.50 (8)	.07	.37	1.81	.97	.98	.053	.000;.096	.40	.0447
RN Valid.	126	1.67 (2)	.01	.01	5.33	.87	.63	.186	.000;.088	.30	.0634

5.1.3.2 Faktor Absolutes Wissen

Das Gesamtmodell mit allen Items führt zu einer sehr schlechten Modellpassung (siehe Tabelle 72). Lediglich der RMSEA und der SRMR entsprechen den Beurteilungskriterien für eine akzeptable Modellpassung. Auf eine Inspektion der Residuen bzw. der Modifikationsindizes auf Itemebene wurde an dieser Stelle verzichtet. Da die Skala *Absolutes Wissen*, gestützt aus den Ergebnissen von Hauptkomponentenanalysen von Moscher, Gruber und der Studienstiftungsgruppe EPI (2005), aus den zwei Faktoren Umgang mit Autoritäten und Sicherheit des Wissens zusammengesetzt wurde, wurde zunächst überprüft, ob diese Zusammenfassung gerechtfertigt war.

⁴⁸ Dies gilt schon bei τ -Äquivalenz, aber jede Skala, welche die Bedingungen für Parallelität erfüllt, erfüllt auch die Bedingungen der τ -Äquivalenz.

Dazu wurde eine EFA nach der Maximum-Likelihood-Methode mit anschließender Promax-Rotation ($\kappa = 2$) durchgeführt. Die Maximum-Likelihood-Methode zur Extraktion der Faktoren wurde aus Vergleichsgründen mit dem Schätzalgorithmus für weiterführende konfirmatorische Faktorenanalysen gewählt. Das KMO-Maß beträgt .76, die Daten sind für die Durchführung einer EFA geeignet, wenn auch das KMO-Maß unter dem von Backhaus et al. (2003) zu fordernden Wert von mindestens .80 liegt. Der Scree-Plot in Abbildung 23 legt eine zweifaktorielle Lösung nahe.

Scree-Plot

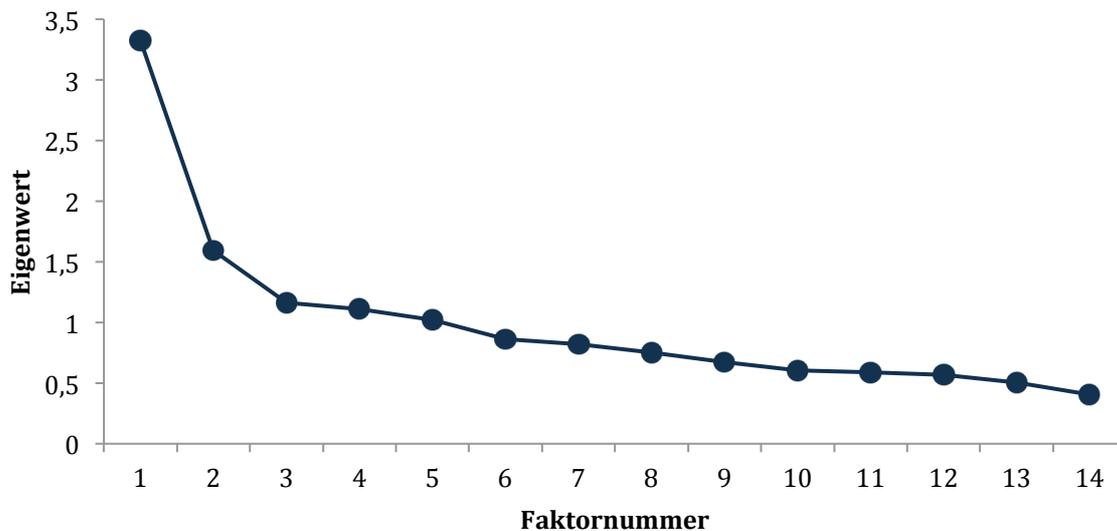


Abbildung 23: Scree-Plot der Skala Absolutes Wissen

Die Zusammenfassung der beiden ursprünglichen Skalen Umgang mit Autoritäten und Sicherheit des Wissens ist dem ersten Anschein nach nicht zu rechtfertigen. Der Skala Sicherheit des Wissens waren ursprünglich die Items 1, 4, 13, 15, 17, 34, 37, 44, 45 und 49 zugeordnet, der Skala Umgang mit Autoritäten die Items 9, 16, 29 und 42. Die Mustermatrix der Hauptachsenanalyse ergibt die in Tabelle 68 dargestellte Zuordnung der Items zu den beiden Faktoren.

Tabelle 69: Mustermatrix der zweifaktoriellen Lösung

Item	Faktor		Ursprüngliche Zuordnung
	1	2	
Item 15	.65	-.13	Sicherheit des Wissens
Item 01	.61	-.02	Sicherheit des Wissens
Item 29	.60	-.05	Umgang mit Autoritäten
Item 04	.52	-.01	Sicherheit des Wissens
Item 09	.51	.09	Umgang mit Autoritäten
Item 17	.31	.16	Sicherheit des Wissens
Item 42	.29	.23	Umgang mit Autoritäten
Item 49	.26	.13	Sicherheit des Wissens
Item 45	-.14	.55	Sicherheit des Wissens
Item 13	-.02	.52	Sicherheit des Wissens
Item 37	.10	.48	Sicherheit des Wissens
Item 34	-.03	.47	Sicherheit des Wissens
Item 16	.08	.40	Umgang mit Autoritäten
Item 44	.08	.35	Sicherheit des Wissens

Die ursprüngliche Zuordnung der Items zu den beiden Skalen Sicherheit des Wissens und Umgang mit Autoritäten spiegelt sich nicht in den Ergebnissen der EFA. In der Mustermatrix zeigt sich, dass die Items 42 und 49 nur gering ($< .30$) laden und Item 42 eine fast gleich große Nebenladung auf dem zweiten Faktor aufweist. Das Item 42 wird somit, wenn auch nur gering, durch beide Faktoren beeinflusst. Da das Ziel die Konstruktion eindimensionaler Messmodelle ist, wird dieses Item aus den nachfolgenden Analysen ausgeschlossen. Der Ausschluss ist auch inhaltlich nicht unbegründet, da sich der Wortlaut des Items eher auf eine Eigenschaft bezieht, die Dozenten im Alltagsdenken zugeschrieben wird. Das Item 49 wird ebenso ausgeschlossen, da es mit keinem der beiden Faktoren assoziiert zu sein scheint. Auch dies ist vom Inhalt des Items her nicht unbegründet, da es sich auf die Charakteristika der wissenschaftlichen Forschung bezieht. Der semantische Gehalt der beiden ausgeschlossenen Items hat keinen Bezug zu den Einstellungen einer Person über Wissen. Somit werden dem Faktor 1 die Items 15, 01, 29, 04, 09, und 17, dem Faktor 2 die Items 45, 13, 37, 34, 16 und 44 zugeordnet.

Der erste Faktor beinhaltet Items, die sich auf die Wahrheit und Eindeutigkeit von wissenschaftlichen Aussagen beziehen. Der zweite Faktor bezieht sich auf die Dauerhaftigkeit und Objektivität von Wissen. Dieses Modell wird nun durch eine zweifaktorielle CFA (Modell 2FM) auf Modellpassung überprüft. Dabei werden der ersten latenten Variable alle Items des ersten Faktors und der zweiten latenten Variable alle Items des zweiten Faktors zugeordnet, es wurden keine Kreuzladungen zugelassen, die beiden latenten Variablen wurden als korrelierend modelliert.

Das ursprüngliche Modell 2FM (vgl. Tabelle 72) ergab kein passendes Messmodell. Eine Untersuchung der Residuen ergab, dass das größte standardisierte Residuum durch die Items 04 und 15 verursacht wird, ebenfalls deutet der Modifikationsindex auf einen korrelierenden Fehler zwischen Item 04 und 15 hin. Item 04 wies eine etwas geringere Faktorladung als Item 15 auf. Inhaltlich besitzt Item 04 mindestens zwei Schwachpunkte. Es ist sowohl unklar, was unter einem Experten zu verstehen ist, als auch, was „Grundverständnis“ bedeutet. Aus diesen Gründen wird Item 04 ausgeschlossen.

Das hieraus resultierende Modell 2FM 1 zeigte wiederum keine ausreichende Modellpassung. Die Items 37 und 34 wiesen ein über dem kritischen Wert von 1.96 liegendes Residuum auf und auch hier deutet der Modifikationsindex auf einen möglicherweise korrelierenden Fehler hin. Da beide Items das Wort objektiv enthalten und inhaltlich den gleichen Bezug haben, ist ein korrelierender Fehler aufgrund einer ähnlichen Itemformulierung nicht ausgeschlossen. Es wurde entschieden, Item 37 wegen der Formulierungen auszuschließen, obwohl es eine höhere Faktorladung als Item 34 aufwies. Der Grund ist die unscharfe Formulierung der Kernaussagen in Item 37. Das Modell nach dem Ausschluss ist 2FM 2.

Die Modellpassung von 2FM 2 erwies sich als sehr gut. Alle Indizes lagen im Bereich guter bis akzeptabler Modellpassung. Eigentlich war keine weitere Modifikation mehr nötig, aber eine Inspektion der Residuen und Modifikationsindizes ergab, dass eine Verbesserung durch Ausschluss von Item 16 erreicht werden kann. Da sich dieses Item auf die Konsultation auf Experten bezieht, wurde es ausgeschlossen.

Das finale Modell 2FM final lässt sich wie folgt zusammenfassen: Der Faktor 1 besteht aus den Items 15, 01, 29, 09 und 17. Inhaltlich bezieht sich der erste Faktor auf eine Absolutheit von Wissen, er beinhaltet die Items, die sich auf die Existenz von Wahrheit bzw. auf die Möglichkeit, wahre Aussagen zu erhalten, beziehen. Der Faktor wird mit *Absolutheit des Wissens* (AW) bezeichnet. Der zweite Faktor besteht den Items 45, 13, 34 und 44 und bezieht sich auf die Existenz und Objektivität wissenschaftlicher Aussagen sowie auf die Dauerhaftigkeit von Erkenntnissen. Dieser Faktor wird mit *Sicherheit/Objektivität des Wissens* (S/OW) bezeichnet. Werden die beiden Modelle AW und S/OW jeweils separat als einzelnes Messmodell geprüft, zeigt jedes für sich ebenfalls eine gute Modellpassung.

Der Mardia-Test ergab eine Abweichung der Daten von der multivariaten Normalverteilung ($M = 13.60$, $CR = 7.07$). Damit ist bei der Interpretation der Parameter das Risiko erhöhter Signifikanzen zu beachten. Die Korrelation der beiden Faktoren beträgt .32 und ist signifikant (vgl. Tabelle 70). Die Faktorladungen sind ebenfalls alle signifikant und bis auf die Ladung des Items 34 substantiell.

Tabelle 70: Parameterwerte für das Modell S/OWF final

Faktor/Item	Partielles Regressionsgewicht		SE	CR	<i>p</i>
	nicht standardisiert	standardisiert			
AW/Item 15	1.00	.45			
AW/Item 01	1.33	.65	.27	4.91	***
AW/Item 29	1.26	.63	.26	4.89	***
AW/Item 09	1.16	.57	.25	4.73	***
AW/Item 17	.77	.33	.22	3.44	***
S/OW/Item 45	1.00	.46			
S/OW/Item 13	1.56	.64	.38	4.07	***
S/OW/Item 34	.63	.29	.22	2.90	***
S/OW/Item 44	1.60	.56	.39	4.12	***
Korrelation AW-S/OW	.10	.32	.04	2.43	.02

Anmerkung: ***: $p < .001$

Die Validierung des ermittelten Modells an der Validierungsstichprobe lieferte allerdings ein uneinheitliches Bild. Für das Gesamtmodell mit zwei Faktoren ergeben sich Fitindizes, die auf eine akzeptable Modellpassung hindeuten. Der TLI lag allerdings außerhalb der Grenzen für eine akzeptable Modellpassung. Um zu ermitteln, ob ggf. beide Faktoren eine nicht akzeptable Modellpassung zeigen oder ob die Ursache der mangelhaften Modellpassung nur in einem der beiden Faktoren zu suchen ist, wurden jeweils beide Faktoren getrennt mit den Daten der Validierungsstichprobe geprüft.

Für den Faktor *Absolutes Wissen* ergaben sich Fitindizes, die alle auf eine gute Modellpassung hindeuten. Für den Faktor *Sicherheit/Objektivität des Wissens* allerdings ergab die Berechnung mithilfe der Validierungsstichprobe Fitindizes, die außerhalb des Bereichs akzeptabler Modellpassung lagen, insbesondere war dies beim RMSEA der Fall. An dieser Stelle kann also festgehalten werden, dass nur der Faktor *Absolutheit des Wissens* sich in der Validierungsstichprobe replizieren lässt. Der zweite Faktor *Sicherheit/Objektivität des Wissens* ist vorläufig nur im Rahmen der Analysestichprobe gültig und kann daher nicht als allgemeingültig angesehen werden. Analog zum Faktor *Reflexive Natur des Wissens* wird der Faktor *Sicherheit/Objektivität des Wissens* für die Regression epistemologischer Überzeugungen auf

Lernstrategien genutzt, ebenfalls mit der Einschränkung, dass die Ergebnisse nur für die vorliegende Stichprobe gelten.

Die Aufteilung in zwei Faktoren wirft gleichzeitig Fragen der diskriminanten Validität auf. Im Hinblick auf die diskriminante Validität geht es um die Frage nach der Höhe des Zusammenhangs der beiden Faktoren *Absolutheit des Wissens* und *Sicherheit/Objektivität des Wissens*. Ist der Zusammenhang hoch, so ist es nicht plausibel, von zwei getrennten Faktoren auszugehen, sodass ggf. anzunehmen ist, dass es sich bei den beiden Faktoren um dasselbe Konstrukt handelt. Allerdings ist die Grenze, die eine hohen von einem niedrigen Zusammenhang trennt, arbiträr (vgl. Borsboom et al., 2004). Ein Ansatz zur Prüfung, ob diskriminante Validität vorliegt oder nicht, ist die Betrachtung der Korrelation zwischen den beiden Faktoren. Diese ist .32 (vgl. Tabelle 71) und mit $p = .02$ signifikant auf dem 5%-Niveau. Das Problem ist aber hier die Festlegung der Grenze, ab welcher eine Korrelation darauf hindeutet, dass zwei Faktoren daselbe Konstrukt erfassen. Eine Korrelation von .32 bedeutet eine Varianzaufklärung von 10.39 %. Die Höhe der Varianzaufklärung ist allerdings zu gering, um anzunehmen, dass die Faktoren ein identisches Konstrukt erfassen. Zu dieser Problematik kommt hinzu, dass zur Prüfung der diskriminanten Validität mithilfe der Korrelation der latenten Variablen eine Versuchsplanung mit der Angabe der maximal zu tolerierenden Korrelation zwischen den beiden Faktoren notwendig wäre. Dies ist erforderlich, um sowohl den α - als auch den β -Fehler zu kontrollieren (Hartig et al., 2008). Eine andere Möglichkeit zur Prüfung der diskriminanten Validität besteht darin, die Korrelation der beiden Faktoren im CFA-Modell gleich eins zu setzen und mithilfe eines χ^2 -Differenztests zu bestimmen, ob diese Restriktion zu einer signifikanten Verschlechterung der Modellpassung führt (Hildebrandt & Temme, 2006). Das Modell mit dieser Restriktion führt zu $\chi^2 = 119.17$ ($df = 27, p < .01$). Damit ergibt sich ein $\Delta\chi^2 = 91.01$, was weit über dem kritischen Wert von 3.84 für einen signifikanten $\Delta\chi^2$ -Wert liegt. Die Restriktion führt also zu einer signifikanten Verschlechterung der Modellpassung, was dafür spricht, dass es sich bei den beiden Faktoren um unterschiedliche Konstrukte handelt. Allerdings muss auch in Bezug auf die Überprüfung der diskriminanten Validität der beiden Faktoren beachtet werden, dass die Ergebnisse nur für die Analytestichprobe Gültigkeit besitzen. Eine Verallgemeinerung ist wegen der fraglichen Modellpassung des Faktors *Sicherheit/Objektivität des Wissens* in der Validierungsstichprobe nicht möglich.

Bzgl. der psychometrischen Eigenschaften muss für ein zweifaktorielles CFA-Modell eine Anpassung des in Kapitel 3.4.3 beschriebenen Verfahrens zur Schätzung der Äquivalenz

durchgeführt werden (Brown, 2006). Dazu werden die Restriktionen immer zuerst für den ersten Faktor eingeführt und dann die Modellpassung berechnet. Anschließend werden die geforderten Restriktionen auch auf dem zweiten Faktor eingeführt. Es ergeben sich also insgesamt vier Modelle, die geprüft werden müssen. Die Berechnung der Skalenreliabilität ρ erfolgt dann mittels des vorgestellten Verfahrens, allerdings mit der Anpassung, dass für jede der beiden latenten Variablen eine Hilfsvariable definiert wird. Der Modellfit des Modells 2FM tau 1, bei dem die Restriktionen der τ -Äquivalenz auf dem ersten Faktor eingeführt wurden, zeigt in allen Fitindizes eine gute Modellpassung. Bei dem Modell 2FM tau 2 liegen alle Fitindizes im Bereich akzeptabler Modellanpassung. Das τ -äquivalente Modell kann als akzeptabel betrachtet werden. Bei dem parallelen Modell 2FM parallel 1 liegen zwar noch einige Fitindizes im Bereich guter bis akzeptabler Modellpassung, allerdings liegt der CFI deutlich unter der Grenze akzeptabler Modellpassung. Beim Modell 2FM parallel 2 sinkt der CFI auf .77 und der nach Bollen-Stine korrigierte p-Wert liegt deutlich unter den Grenzwerten für eine akzeptable Modellpassung. Bei dem zweifaktoriellen CFA-Modell kann also höchstens die Gültigkeit der τ -Äquivalenz nachgewiesen werden. Die Reliabilitätskoeffizienten sind in Tabelle 71 zusammengefasst.

Tabelle 71: Reliabilitätsschätzungen der Skalen für Modell S/OWF final

Reliabilitäten
Skalenreliabilität ρ AW kongenerisch = .67
Skalenreliabilität ρ S/OW kongenerisch = .56
Skalenreliabilität ρ AW tau = .67
Skalenreliabilität ρ S/OW tau = .53
Cronbach's α AW = .67
Cronbach's α S/OW = .53

Die Reliabilitätskoeffizienten liegen auch hier im Bereich, wie sie bei Skalen zu epistemologischen Überzeugungen üblich sind. Allerdings hat insbesondere der Faktor *Sicherheit/Objektivität* des Wissens eine sehr niedrige Reliabilität. Die Items scheinen also eine Dimension interindividueller Unterschiede in Bezug auf epistemologische Überzeugungen zu erfassen, allerdings sind die Items nicht für eine reliable Erfassung dieser Dimension geeignet.

Zusammenfassend lässt sich feststellen, dass die ursprüngliche Zusammenlegung von zwei Dimensionen bei der durch eine konfirmatorische Faktorenanalyse erfolgende Prüfung nicht zu rechtfertigen war. Bei der Anwendung einer Hauptachsenanalyse ergab sich eine klare zweidimensionale Struktur und auch die Ergebnisse der konfirmatorischen Faktorenanalyse

Ergebnisse zum EPI

stützen die Annahme von zwei Faktoren, wobei aber die Kreuzvalidierung des Faktors *Sicherheit/Objektivität des Wissens* fraglich ist. Problematisch hingegen sind die niedrigen Reliabilitäten, vor allem die der Skala zur Erfassung des Faktors *Sicherheit/Objektivität des Wissens*.

Tabelle 72: Ergebnisse des zweifaktoriellen CFA-Modells

Modell	N	χ^2 (df)	p	Bollen-Stine-p	χ^2/df	CFI	TLI	RMSEA	CI (90 %)	p_{close}	SRMR
AW urspr.	291	246.13 (77)	***	.01	3.20	.70	.65	.087	.075;.099	***	.0775
2FM	291	127.40 (53)	***	.01	2.40	.85	.80	.070	.054;.085	.02	.0608
2FM 1	291	85.64 (43)	***	.02	1.99	.89	.86	.058	.040;.076	.21	.0551
2FM 2	291	46.13 (34)	.08	.29	1.36	.96	.95	.035	0;.059	.84	.0476
2FM final	291	28.16 (26)	.35	.57	1.08	.99	.99	.017	0;.051	.95	.0399
AW	291	5.61 (5)	.35	.46	1.12	.99	.88	.021	0;.086	.69	.0263
S/OW	291	3.85 (2)	.15	.26	1.93	.98	.93	.057	0;.142	.34	.0288
2FM S1	207	32.81	.17	.35	1.26	.96	.95	.036	0;.069	.72	.0527
2FM S2	221	21.21	.73	.85	.82	1.00	1.03	.000	0;.040	.98	.0374
2FM S3	217	23.89	.58	.78	.92	1.00	1.02	.000	0;.048	.96	.0409
2FM S4	212	28.72	.32	.53	1.10	.99	.98	.022	0;.060	.86	.0510
2FM S5	221	21.30	.73	.84	.820	1.00	1.03	.000	0;.040	.98	.0406
2FM S6	215	19.60	.81	.88	.76	1.00	1.04	.000	0;.035	.99	.0399
2FM S7	217	21.87	.70	.78	.84	1.00	1.03	.000	0;.043	.98	.0416
2FM S8	221	23.77	.58	.77	.91	1.00	1.02	.000	0;.048	.96	.0421
2FM S9	219	41.01	.03	.12	1.58	.93	.90	.052	.016;.080	.43	.0542
2FM S10	214	27.67	.38	.54	1.06	.99	.99	.017	0;.0058	.89	.0476
2FM kongen	291	28.16 (26)	.35	.57	1.08	.99	.99	.017	0;.051	.95	.0399
2FM tau 1	291	33.98 (30)	.28	.51	1.13	.99	.98	.021	0;.051	.94	.0438
2FM tau 2	291	53.29 (33)	.01	.08	1.62	.93	.92	.046	.021;.068	.59	.0574
2FM para. 1	291	73.25 (37)	***	.01	1.98	.87	.88	.058	.038;.078	.23	.0622
2FM para. 2	291	104.52 (40)	***	***	2.61	.77	.79	.075	.057;.092	.01	.0578
2FM Val.	126	49.30 (26)	***	.08	1.90	.91	.87	.085	.047;.120	.06	.0719
AW Val.	126	3.31 (5)	.65	.72	.66	1.00	1.03	.000	0;1.000	.78	.0267
S/OW Val.	126	6.21 (2)	.05	.13	3.10	.96	.87	.130	.017;.252	.09	.0478

Anmerkung: ***: $p < .001$

5.1.3.3 Faktor Soziale Komponente des Wissens

Das initiale Modell erreichte keine zufriedenstellende Modellanpassung. Fast alle Fitindizes deuten auf eine nicht mehr akzeptable Modellpassung hin. Die Inspektion der Residuen deutete ebenso wie der Modifikationsindex auf einen nicht modellierten korrelierenden Messfehler der Items 20 und 41 hin. Bei Betrachtung des jeweiligen Inhalts der beiden Items erscheint ein korrelierender Messfehler nicht plausibel. Auffällig ist, dass in beiden Items das Wort „nur“ enthalten ist, allerdings ist fraglich, ob dies ausreicht, um eine Korrelation der Fehlervariablen zu bedingen. Das Item 41 hat jedoch von allen Items die kleinste Faktorladung und bezieht sich auf das Ausmaß der sozialen Komponente des Wissens, was im Extremfall bedeuten kann, dass nur solches Wissen als gültig angesehen wird, das mit anderen Personen geteilt wird. Dabei wird aber auch die Fähigkeit einer Person zur Wissensvermittlung angesprochen. Im Folgenden wurde Item 41 aus der weiteren Analyse ausgeschlossen, das resultierende Modell ist SK 1 (vgl. Tabelle 75).

Das Modell SK 1 weist immer noch eine nicht akzeptable Modellanpassung auf. Die Inspektion der Residuen und der Modifikationsindizes legten die Korrelation der Fehlervariablen zwischen den Items 20 und 40 nahe. Item 40 und Item 20 zielen inhaltlich auf den gleichen Sachverhalt, nämlich die Existenz von Wissen unabhängig von der Verbreitung dieses Wissens (lediglich die Polung der Items ist entgegengesetzt). Da die Formulierung von Item 40 gegenüber der Formulierung von Item 20 weniger genau ist und sich sowohl auf die Geheimhaltung von Wissen durch eine Person als auch auf die Nichtveröffentlichung von gewonnenen Erkenntnissen bezieht, wurde dieses Item aus dem Messmodell entfernt.

Das resultierende Modell SK final zeigte in allen Fitindizes eine gute Modellpassung. Ebenso führt das Modell in der Validierungsstichprobe zu einer guten Modellpassung. Das Messmodell scheint somit keine Anpassung an Besonderheiten der Analysestichprobe wiederzugeben.

Die Daten wiesen eine Abweichung von der multivariaten Normalverteilung ($M = 11.31$, $CR = 11.94$) auf. Die Schiefe und die Wölbung der einzelnen Variablen lagen aber innerhalb der Grenzen von West et al. (1995). Allerdings muss bei der Interpretation der Parameter mit überhöhten Signifikanzen gerechnet werden. Die Parameter sind in Tabelle 73 dargestellt.

Tabelle 73: Parameter der Skala Soziale Komponente des Wissens

Item	partielles Regressionsgewicht		SE	CR	<i>p</i>
	nicht standardisiert	standardisiert			
Item 20	1.00	.33			
Item 22	-1.42	-.34	.53	-2.66	.01
Item 25	-1.83	-.65	.62	-2.93	***
Item 52	1.17	.53	.38	3.06	***

Anmerkung: ***: $p < .001$

Das Messmodell mit den Restriktionen τ -Äquivalenz führte zu einer akzeptablen Modellpassung, das parallele Messmodell erzielte dagegen keine akzeptable Modellpassung mehr. In Tabelle 74 finden sich die Reliabilitätsschätzungen für das finale Modell.

Tabelle 74: Reliabilitätsschätzungen der Skala Soziale Komponente des Wissens

Reliabilitäten
Skalenreliabilität ρ kongenerisch = .50
Skalenreliabilität ρ tau = .49
Cronbachs α = .49

Die Reliabilitäten fallen sehr niedrig aus. Auch bei diesem Faktor scheint eine Dimension interindividueller Unterschiede erfasst zu werden, allerdings nicht sehr zuverlässig. Das Messmodell erfüllt die Bedingungen der τ -Äquivalenz, die Werte der Skalenreliabilität ρ und Cronbachs α sind in etwa auch gleich hoch. Cronbachs α liegt mit .49 aber weit unterhalb des akzeptablen Grenzwertes. Analog zu den Überlegungen bei einigen Faktoren des Schommer-Fragebogens (vgl. Kapitel 1.3.1) kann es möglich sein, dass diesem Faktor zwar eine gemeinsame Dimension zugrunde liegt, allerdings besteht die Möglichkeit, dass es sich hierbei um ein *Higher-order*-Faktormodell bzw. ein Modell mit hierarchischen Faktoren handeln könnte (Yung, Thissen & McLeod, 1999). Die Präzision von Cronbachs α liegt bei .0346 und damit in einem Bereich, in dem von Multidimensionalität ausgegangen werden kann. Inhaltlich ist diese Möglichkeit auch plausibel, da sich die Items auf unterschiedliche Aspekte beziehen. So bezieht sich Item 20 darauf, dass Wissen nur dann als gültig angesehen werden kann, wenn es veröffentlicht wurde. Die Items 22 und 25 hingegen beziehen sich darauf, dass es Wissensinhalte gibt, die nur einem Individuum bekannt sind. Item 52 beinhaltet, dass einer Person nur dann Wissen zugeschrieben werden kann, wenn sie dies in Interaktion mit anderen zeigt. Die Items lassen sich inhaltlich somit drei Kategorien zuordnen, wobei sich aber alle Items auf die Vorstellung beziehen, dass Wissen nur dann gültig ist, wenn es mit anderen geteilt wird. Die Items gestatten somit die Erfassung des Faktors Soziale Komponente des Wissens. Das Mo-

dell erzielte einen guten Modellfit in der Validierungsstichprobe. Allerdings gelingt die Erfassung des Konstrukts nicht reliabel.

Tabelle 75: Ergebnisse für die Skala Soziale Komponente des Wissens

Modell	N	χ^2 (df)	p	Bollen-Stine- p	χ^2/df	CFI	TLI	RMSEA	CI (90 %)	p_{close}	SRM R
SK	291	38.81 (9)	***	***	4.31	.79	.64	.107	.074;.142	.01	.0651
SK 1.	291	2.06 (5)	***	.02	4.01	.86	.71	.102	.058;.150	.03	.0538
SK final	291	.95 (2)	.62	.76	.48	1.00	1.04	.000	0;.094	.79	.0148
SK S1	207	1.20	.53	.65	.64	1.00	1.04	.000	0;.121	.67	.0203
SK S2	221	3.78	.15	.29	1.89	.97	.90	.064	0;.162	.30	.0346
SK S3	217	.30	.86	.89	.15	1.00	1.10	.000	0;.071	.92	.0098
SK S4	212	.44	.80	.86	.22	1.00	1.10	.000	0;.086	.87	.0117
SK S5	221	1.61	.44	.59	.80	1.00	1.02	.000	0;.125	.62	.0218
SK S6	215	2.09	.35	.50	1.04	1.00	1.00	.014	0;.137	.54	.0272
SK S7	217	.49	.78	.85	.24	1.00	1.10	.000	0;.087	.87	.0130
SK S8	221	1.44	.49	.60	.72	1.00	1.03	.000	0;.121	.65	.0204
SK S9	219	3.11	.21	.37	1.55	.98	.94	.050	0;.153	.38	.0310
SK S10	214	1.81	.40	.59	.91	1.00	1.01	.000	0;.132	.57	.0235
SK kong.	291	.95 (2)	.62	.76	.48	1.00	1.04	.000	0;.094	.79	.0148
SK tau	291	9.06 (5)	.11	.31	1.81	.94	.93	.053	0;.107	.40	.0488
SK paral- lel	291	139.68 (8)	***	***	17.46	0.00	-.32	.238	.204;.274	***	.0606
SK Val.	126	.07 (2)	.94	.96	.04	1.00	1.30	.000	.000;.000	.97	.0062

Anmerkung: ***: $p < .001$

5.1.3.4 Faktor Value of Knowledge

Das ursprüngliche Modell für den Faktor *Value of Knowledge* bestand aus den Items 12, 21, 24, 38 und 53. Dieses Modell ergab außer beim SRMR keine ausreichende Modellpassung. Die Inspektion der standardisierten Residuen zeigte, dass durch das Modell die Kovarianz von Item 21 und 53 erheblich unterschätzt wird. Der Modifikationsindex deutet auf einen korrelierenden Fehler der beiden Items hin, was durch den Inhalt der Items gestützt wird. Beide Items beziehen sich auf den Nutzen von Wissen, wobei Item 21 die Relevanz von Wissen für praktische Fragestellungen betont, während Item 53 den gesellschaftlichen Nutzen von Wissen thematisiert. Item 21 wies von allen Items die niedrigste Faktorladung auf. Die empirischen Befunde legten den Ausschluss von Item 21 nahe. Das Modell ohne Item 21 (VK final) zeigte in allen berechneten Fitindizes eine gute Modellpassung, allerdings ist auch bei diesem Modell das Konfidenzintervall des RMSEA sehr breit. In der Validierungsstichprobe ergab sich eine gute Modellpassung, sodass nicht davon ausgegangen werden kann, dass das Messmodell an die Besonderheiten der Analytestichprobe angepasst wurde.

Die Variablen des Modells weichen nicht von der multivariaten Normalverteilung ab, der Mardia-Test ist nicht signifikant ($M = -.05$, $CR = -.06$).

Tabelle 76: Parameterwerte des Faktors Value of Knowledge

Item	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
Item 12	1.00	.36			
Item 24	2.89	.77	.75	3.87	***
Item 38	2.26	.61	.57	4.00	***
Item 53	1.55	.44	.44	3.57	***

Anmerkung: ***: $p < .001$

Jedes Item weist eine substanzielle und signifikante Faktorladung auf. Das kongenerische Modell ohne jegliche Modellrestriktion passt gut mit den empirischen Daten zusammen, das τ -äquivalente und parallele Modell passen jedoch überhaupt nicht zur Datenstruktur. Die geschätzten Reliabilitäten sind in Tabelle 77 dargestellt, diese fallen auch bei hier niedrig aus. Wegen der fehlenden τ -Äquivalenz stellt Cronbachs α hier keine sinnvolle Schätzung der Reliabilität dar. Da das Messmodell keine korrelierenden Fehler enthält und die Bedingung der τ -Äquivalenz nicht zutrifft, unterschätzt Cronbachs α die Reliabilität der Skala (Raykov, 2001b, Zimmermann, 1972) und die Skalenreliabilität ρ stellt den passenden Schätzer dar. Der numerische Wert für Cronbachs α liegt auch erheblich unter dem Wert für Skalenreliabilität ρ für das kongenerische Messmodell. Allerdings ist die Skalenreliabilität als solche auch unterhalb des akzeptablen Grenzwertes von .70.

Tabelle 77: Reliabilitätsschätzungen für die Skala Value of Knowledge

Reliabilitäten
Skalenreliabilität ρ kongenerisch = .64
Cronbachs α = .60

Die Items gestatten somit die Erfassung der latenten Variablen *Value of Knowledge*, allerdings ist auch hier unreliabel. Ein Grund hierfür könnte darin zu sehen sein, dass die Items sich einmal auf die Bedeutung der praktischen Relevanz von Wissen und einmal auf den Nutzen, der Wissen für die Gesellschaft hat, beziehen. Auch hier könnte der Bezug auf verschiedene Aspekte in den Items einen Grund für die mangelnde Reliabilität der Items darstellen.

Tabelle 78: Fitindizes des Faktors Value of Knowledge

Modell	N	χ^2 (df)	p	Bollen-Stine- p	χ^2/df	CFI	TLI	RMSEA	CI (90 %)	p_{close}	SRMR
VK	291	29.62 (5)	***	***	5.92	.85	.71	.130	.087;.177	.01	.0647
VK final	291	3.53 (2)	.17	.23	1.77	.99	.96	.051	0;.138	.38	.0238
VK S1	207	5.40	.07	.06	2.70	.97	.91	.091	0;.187	.17	.0351
VK S2	221	4.78	.09	.11	2.39	.97	.91	.080	0;.174	.22	.0322
VK S3	217	1.52	.47	.51	.76	1.00	1.02	.000	0;.124	.63	.0185
VK S4	212	2.43	.23	.31	1.22	.99	.98	.032	0;.144	.47	.0240
VK S5	221	3.27	.20	.24	1.63	.99	.96	.054	0;.155	.36	.0263
VK S6	215	3.20	.20	.19	1.60	.99	.96	.053	0;.156	.37	.0266
VK S7	217	1.73	.42	.48	.87	1.00	1.01	.000	0;.129	.59	.0213
VK S8	221	2.26	.32	.34	1.13	1.00	.99	.024	0;.138	.50	.0229
VK S9	219	5.64	.06	.07	2.82	.96	.88	.091	0;.185	.16	.0342
VK S10	214	3.53	.17	.20	1.77	.99	.95	.060	0;.161	.33	.0282
VK kon- gen.	291	3.53 (2)	.17	.23	1.77	.99	.96	.051	0;.138	.38	.0238
VK tau	291	39.74 (5)	***	***	7.95	.73	.67	.155	.112;.201	***	.0963
VK paral- lel	291	51.34 (8)	***	***	6.42	.66	.74	.137	.102;.173	***	.0826
VK Val.	126	.53 (2)	.77	.83	.27	1.00	1.04	.000	.000;.118	.82	.0130

Anmerkung: ***: $p < .001$

5.1.3.5 Faktor Geschlechtsspezifische Wissenszugänge

Das ursprüngliche Modell, das für den Faktor *Geschlechtsspezifische Wissenszugänge* postuliert wurde, wies eine sehr schlechte Modellpassung auf ($\chi^2 = 412.99$ ($df = 35$), $p < .001$, Bollen-Stine- $p = .001$, $\chi^2/df = 11.80$, CFI = .66, TLI = .56, SRMR = .1287, RMSEA = .193, CI = [.177;.210], $p_{close} < .001$). Die Inspektion der Residuen ergab, dass viele Korrelationen durch das Modell nicht richtig repliziert wurden. Da die Residuen und die Modifikationsindizes keinen Hinweis auf eine mögliche Fehlspezifikation gaben, wurde eine EFA durchgeführt. Analog dem Vorgehen bei dem Faktor Absolutes Wissens wurde eine Maximum-Likelihood-Faktorenanalyse mit Promax-Rotation ($\kappa = 4$, KMO = .83) durchgeführt. Der Scree-Plot deutet auf eine zweifaktorielle Lösung hin.

Die Mustermatrix weist in der zweifaktoriellen Lösung eine klare Struktur auf, lediglich Item 8 lädt auf beiden Faktoren, allerdings sind die Faktorladungen auf beiden Faktoren nur knapp über .30.

Scree-Plot

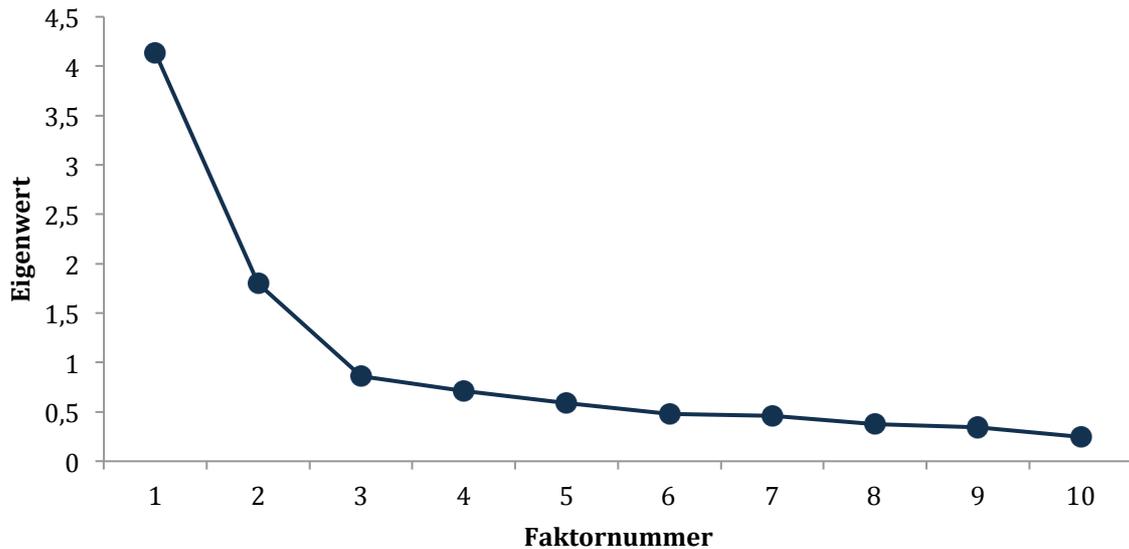


Abbildung 24: Scree-Plot der Faktorenanalyse der Skala Geschlechtsspezifische Wissenszugänge

Tabelle 79: Mustermatrix der zweifaktoriellen Lösung

Item	Faktor	
	1	2
Item 36	.84	.12
Item 46	.77	.03
Item 30	.76	.07
Item 23	.66	-.14
Item 43	.53	-.20
Item 19	.10	.89
Item 05	.02	.84
Item 14	.00	.59
Item 18	-.19	.42
Item 08	.32	-.33

Auf dem ersten Faktor laden die Items 36, 46, 30, 23 und 43. Dies sind alle Items, die auf Geschlechtsunterschiede hinweisen, bis auf Item 30 enthält auch jedes Item das Wort „unterschiedlich“. Auf dem zweiten Faktor laden die Items 19, 05, 14 und 18. Dies sind alle Items, die auf Gleichheit der Geschlechter hinweisen, alle Items enthalten auch das Wort „gleich“. Dies legt den Verdacht nahe, dass es sich bei der zweifaktoriellen Lösung um einen Methodeneffekt handeln könnte, der durch Aufteilung der Items in solche, die sich auf Unterschiede und solche Items, die sich auf die Gleichheit von Geschlechtern beziehen, hervorgerufen wurde. Der Einfluss solcher Formulierungen (insbesondere positiv und negativ formulierter Items) wurde mehrfach nachgewiesen (vgl. zusammenfassend Brown, 2006).

Im Folgenden wird deshalb ein CFA-Modell erstellt, das zwei Methodenfaktoren postuliert und in Abbildung 23 dargestellt ist.

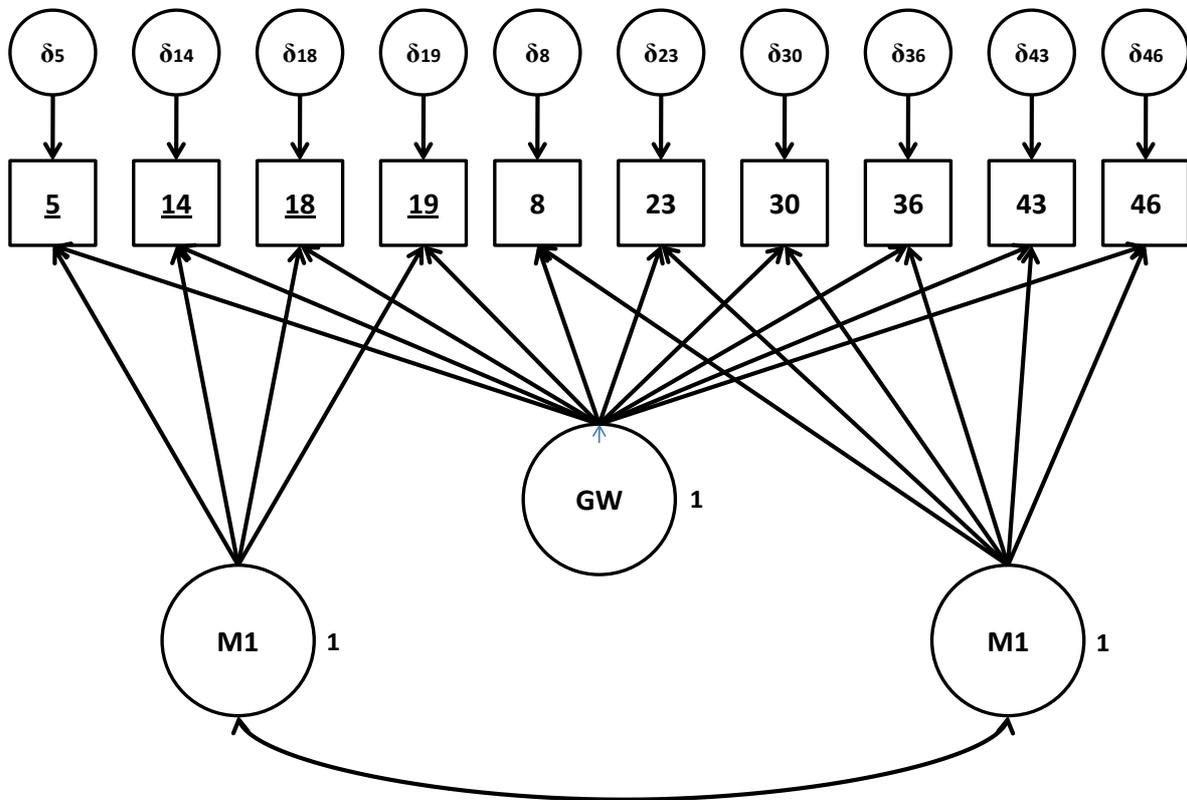


Abbildung 25: CFA-Modell mit zwei Methodenfaktoren für die Skala Geschlechtsspezifische Wissenszugänge (unterstrichene Items haben ein negatives Regressionsgewicht)

Das Modell besagt, dass die manifesten Variablen von der eigentlich interessierenden latenten Variablen Geschlechtsspezifische Wissenszugänge (GW) sowie durch zwei Methodenfaktoren M1 und M2 beeinflusst werden.⁴⁹ Die Varianzen der latenten Variablen dienen der Skalierung und sind auf eins fixiert. Das Modell hat eine akzeptable Modellpassung ($\chi^2 = 68.22$ ($df = 24$), $p < .01$, Bollen-Stine- $p = .002$, $\chi^2/df = 2.84$, CFI = .96, TLI = .92, SRMR = .0384, RMSEA = .080, CI = [.058;.102], $p_{close} = .01$). Allerdings deuten einige der Fitindizes darauf hin, dass es in dem Modell Fehlspezifikationen gibt. Eine Inspektion der Residuen ergab, dass das Modell die empirische Korrelation der Items 14 und 18 sehr stark unterschätzt. Item 14 bezieht sich darauf, dass Männer und Frauen gleich viel wissen und Item 18 beinhaltet, dass Männer und Frauen in den Naturwissenschaften gleich begabt sind. Da diese Items inhaltlich wie auch in der Formulierung nichts gemein haben, ist ein Ausschluss eines dieser Items auf der Grundlage eines empirischen Befundes nicht sinnvoll. Allerdings bezieht sich neben Item

⁴⁹ Alternativ wäre es hier auch möglich gewesen, die Fehlervariablen der Items korrelieren zu lassen, die sich jeweils auf die Unterschiede oder die Gleichheit der Geschlechter beziehen. Dies hätte jedoch zur Folge gehabt, dass das Modell nicht identifizierbar gewesen wäre.

18 auch Item 43 auf die Begabung von Männern und Frauen. Wegen dieses Bezugs auf die Begabung wurden diese beiden Items aus dem CFA-Modell in Abbildung 25 entfernt.

Bei einer erneuten Berechnung des modifizierten Modells ohne die Items 18 und 43 deutete jeder der Fitindizes auf eine gute Modellpassung hin ($\chi^2 = 11.21$ ($df = 11$), $p = .43$, Bollen-Stine- $p = .58$, $\chi^2/df = 1.02$, CFI = 1.00, TLI = 1.00, SRMR = .0170, RMSEA = .008, CI = [.000;.062], $p_{close} = .87$). In der Validierungsstichprobe führte das Modell ebenfalls zu guten Fitindizes ($\chi^2 = 16.37$ ($df = 11$), $p = .13$, Bollen-Stine- $p = .22$, $\chi^2/df = 1.49$, CFI = .98, TLI = .96, SRMR = .0350, RMSEA = .062, CI = [.000;.122], $p_{close} = .33$). Daher kann davon ausgegangen werden, dass die Annahme von Methodenfaktoren allgemein zutrifft und nicht nur eine Besonderheit der Analysestichprobe wiedergibt.

Der Faktor *Geschlechtsspezifische Wissenszugänge* wird in den folgenden Analysen nicht weiter beachtet werden. Die Items beziehen sich nicht direkt auf Geschlechtsunterschiede von epistemologischen Überzeugungen. Vielmehr beinhalten die Items Aussagen, die sich auf interindividuelle Differenzen von Personen in der Einstellung beziehen, dass zwischen den Geschlechtern unterschiedliche Zugänge zu Wissen feststellbar sind. Dies ist aber für eine Betrachtung von Geschlechtsunterschieden epistemologischer Überzeugungen nicht relevant. Relevant für eine solche Betrachtung wäre, ob sich bei Frauen und Männern unterschiedliche epistemologische Überzeugungen nachweisen lassen und nicht ob es Einstellungen zur der Frage gibt, ob Frauen und Männer unterschiedlichen Zugänge zu Wissen aufweisen.

5.1.3.6 Faktor Kulturspezifische Wissenszugänge

Beim Ausgangsmodell lag bis auf dem SRMR keiner der berechneten Fitindizes im Rahmen einer akzeptablen Modellpassung. Eine Inspektion der Residuen ergab, dass das größte Residuum bei der Korrelation von Item 10 und 26 auftrat, ebenso trat bei einer möglichen Korrelation der Fehlervariablen dieser beiden Items der größte Modifikationsindex auf. Item 26 hat überdies die betragsmäßig kleinste Faktorladung. Inhaltlich bezieht es sich nicht explizit auf Kulturunterschiede und die Formulierung hat ihren Schwerpunkt auf Lernen und nicht dem Umgang mit Wissen. Deshalb wird es für die folgenden Analysen ausgeschlossen.

Das resultierende Modell KW 1 zeigte zwar eine akzeptable Modellpassung, die Inspektion der Residuen ergab jedoch, dass der Zusammenhang von Item 11 und 39 überschätzt wird, wobei die Höhe des Residuums nicht außerhalb des kritischen Wertes von 1.96 liegt. Wegen

der inhaltlichen Redundanz und der im Vergleich zu Item 39 niedrigeren Faktorladung wurde Item 11 aus der weiteren Analyse ausgeschlossen. Da Item 11 der Skalierer des Modells war, wurde die Faktorladung von Item 32 zur Skalierung der latenten Variablen auf eins fixiert.

Die Fitindizes des finalen Modells KW final liegen alle im Bereich einer guten Modellpassung. Aber auch hier fällt wie bei einigen anderen Faktoren auf, dass das Konfidenzintervall des RMSEA relativ breit ist. In der Validierungsstichprobe ergaben sich für das Modell allerdings Fitindizes, die außerhalb des Bereichs einer akzeptablen Modellpassung liegen. Man kann bei dem Faktor *Kulturabhängige Wissenszugänge* nicht davon ausgehen, dass das Messmodell sich in anderen Stichproben replizieren lässt.

Die Werte von Schiefe und Wölbung der Variablen liegen alle innerhalb der Grenzen von West et al. (1995), jedoch ist der Mardia-Test auf multivariate Normalverteilung signifikant ($M = 12.18$, $CR = 10.64$). Bei der Interpretation der Parameter in Tabelle 80 ist also mit überhöhten Signifikanzen zu rechnen. Die Faktorladungen sind alle signifikant und auch substantiell.

Tabelle 80: Parameter des Faktors Kulturspezifische Wissenszugänge

Item	partielles Regressionsgewicht		SE	CR	<i>p</i>
	nicht standardisiert	standardisiert			
Item 32	1.00	.48			
Item 35	1.29	.75	.21	6.01	***
Item 39	1.10	.68	.19	5.85	***
Item 10	-.96	-.47	.20	-4.81	***
Item 50	1.17	.65	.20	5.75	***

Anmerkung: ***: $p < .001$

Das Modell erfüllt die Bedingungen der τ -Äquivalenz, aber nicht die der Parallelität. Die Reliabilitätsschätzungen sind in Tabelle 81 zu finden. Die Reliabilität liegt im befriedigenden Bereich.

Tabelle 81: Reliabilitätsschätzungen für die Skala Kulturspezifische Wissenszugänge

Reliabilitäten
Skalenreliabilität ρ kongenerisch = .74
Skalenreliabilität ρ tau = .74
Cronbachs α = .74

Inhaltlich drückt der Faktor individuelle Differenzen im Denken über Kulturunterschiede aus. Ähnlich wie beim Faktor *Geschlechtsspezifische Wissenszugänge* wird mit den Items nicht auf die Erfassung von Kulturunterschieden abgezielt, sondern er wird festgestellt, ob es inter-individuelle Unterschiede im Denken über den Umgang mit Wissensinhalten gibt. Die eigentliche Frage, ob es Kulturunterschiede in den epistemologischen Überzeugungen gibt, wird durch den Nachweis eines solchen Faktors nicht beantwortet. Auch wird, analog zum notwendigen Vorgehen zum Nachweis von Geschlechtsunterschieden, ein Untersuchungsdesign gefordert, welches zum Nachweis von Kulturunterschieden in epistemologischen Überzeugungen geeignet ist. Aus diesem Grund und bedingt durch die Tatsache, dass das Messmodell nicht in einer unabhängigen Stichprobe validiert werden konnte, wird dieser Faktor bei den nachfolgenden Untersuchungen nicht weiter verwendet werden.

Tabelle 82: Fitindizes des Faktors Kulturspezifische Wissenszugänge

Modell	N	χ^2 (df)	p	Bollen-Stine- p	χ^2/df	CFI	TLI	RMSEA	CI (90 %)	p_{close}	SRMR
KW	291	6.648 (14)	***	***	4.33	.89	.84	.107	.080;.136	***	.0637
KW 1	291	2.511 (9)	.02	.12	2.28	.97	.95	.066	.028;.105	.21	.0367
KW final	291	2.917 (5)	.71	.82	.58	1.00	1.01	.000	.000;.061	.91	.0169
KW S1	207	7.874	.16	.31	1.58	.99	.97	.053	.000;.119	.40	.0329
KW S2	221	4.614	.47	.64	.92	1.00	1.00	.000	.000;.090	.73	.0246
KW S3	217	2.649	.75	.86	.53	1.00	1.03	.000	.000;.066	.90	.0192
KW S4	212	3.582	.61	.79	.72	1.00	1.01	.000	.000;.080	.82	.0212
KW S5	221	6.000	.31	.52	1.20	1.00	.99	.030	.000;.0102	.59	.0290
KW S6	215	3.237	.66	.81	.65	1.00	1.01	.000	.000;.075	.85	.0293
KW S7	217	3.630	.60	.76	.73	1.00	1.01	.000	.000;.080	.82	.0220
KW S8	221	3.685	.60	.78	.74	1.00	1.01	.000	.000;.080	.81	.0220
KW S9	219	7.952	.16	.25	1.59	.99	.98	.052	.000;.117	.41	.0300
KW S10	214	3.853	.57	.69	.77	1.00	1.01	.000	.000;.083	.79	.0227
KW kong.	291	2.917 (5)	.71	.82	.58	1.00	1.01	.000	.000;.061	.91	.0169
KW tau	291	14.905 (9)	.10	.30	1.66	.98	.98	.048	.000;.089	.49	.0462
KW paral- lel	291	78.865 (13)	***	***	6.07	.78	.83	.132	.105;.161	***	.0821
KW Val.	126	17.519 (5)	.01	.02	3.50	.92	.83	.142	.073;.216	.02	.0607

Anmerkung: ***: $p < .001$

5.1.3.7 Faktor Lernen lernen

Das initiale Modell, das aus den Items 2, 3, 6, 31, 33 und 47 besteht, zeigte keine akzeptable Modellpassung, lediglich der RMSEA und SRMR lagen knapp im akzeptablen Bereich. Die Inspektion der Residuen ergab, dass die Items 03 und 47 das größte standardisierte Regressionsresiduum aufwiesen. Der Modifikationsindex deutete ebenso auf möglicherweise korrelierende Fehlervariablen der beiden Items hin. Die beiden Items beinhalten die gleiche Aussage, sind nur in der Formulierung entgegengesetzt. Deswegen sollte eines der beiden Items von der weiteren Analyse ausgeschlossen werden. Da Item 03 das kleinere standardisierte Regressionsgewicht aufwies, wurde dieses ausgeschlossen.

Das resultierende Modell LL final besitzt eine gute Modellpassung. In der Validierungsstichprobe ergaben sich gute bis akzeptable Fitindizes, sodass davon ausgegangen werden kann, dass das Messmodell auch auf andere Stichproben als die Analysestichprobe übertragbar ist. Die Schiefe und Wölbung der Variablen liegen in den Grenzen von West et al. (1995), der Mardia-Test lässt eine Abweichung der Daten von der multivariaten Normalverteilung erkennen ($M = 12.25$, $CR = 12.49$). Dies ist bei der Interpretation der Parameterwerte in Tabelle 83 zu berücksichtigen.

Tabelle 83: Parameterwerte für den Faktor Lernen lernen

Item	partiell Regressionengewicht		SE	CR	<i>p</i>
	nicht standardisiert	standardisiert			
Item 02	1.00	.23			
Item 06	-3.60	-.70	1.10	-3.28	***
Item 31	-2.28	-.47	.73	-3.11	***
Item 33	-2.64	-.71	.80	-3.28	***
Item 47	-2.36	-.62	.73	-3.24	***

Anmerkung: ***: $p < .001$

Alle Regressionsgewichte sind signifikant und bis auf das Regressionsgewicht von Item 02 auch substantiell. Der Grund hierfür könnte der inhaltliche Bezug von Item 02 darstellen. Das Item bezieht sich nicht auf die Fähigkeit zum Lernen, sondern auf die zum Lernen erforderliche Begabung. Da das Modell aber eine gute Modellpassung gezeigt hat, wurde zur Vermeidung einer Überanpassung des Modells an die Daten auf einen Ausschluss dieses Items verzichtet.

In Bezug auf die psychometrischen Eigenschaften ist das Modell kongenerisch, allerdings zeigt sich bei den erforderlichen Restriktionen für τ -Äquivalenz und Parallelität keine Passung. Die Reliabilitätsschätzungen sind in Tabelle 84 aufgeführt.

Tabelle 84: Reliabilitätsschätzungen für die Skala Lernen lernen

Reliabilitäten
Skalenreliabilität ρ kongenerisch = .68
Cronbachs α = .66

Die Items haben mangelnde Reliabilität. Der Grund hierfür kann in dem Item 02 gesehen werden, das sich wie oben angemerkt auf Begabung und nicht auf die Fähigkeit zum Erwerb von Lerntechniken bezieht. In der Tat würde die Skalenreliabilität ρ für das kongenerische Modell auf .72 steigen, wenn Item 02 aus der Berechnung ausgeschlossen würde.

Der Faktor bildet ein Konglomerat aus den beiden hypothetischen Konstrukten *Fixed Ability* und *Strategische Komponente des Lernens* (vgl. Kapitel 1.3.3). Aus einer theoretischen Perspektive sind die Konstrukte *Strategische Komponente des Lernens* und *Fixed Ability* zusammenhängend. Die Begabung einer Person beeinflusst den Erwerb bestimmter Lerntechniken, und ebenso wird der Erfolg bestimmter Lerntechniken von der Begabung beeinflusst. An dieser Stelle ist zu überlegen, ob die beiden Konstrukte als *Higher-order*-Faktormodell (bzw. als hierarchisches Faktorenmodell, vgl. Yung et al., 1999) dargestellt werden können. Dabei würden die beiden Konstrukte *Fixed Ability* und *Strategische Komponente des Lernens* First-order-Faktoren bilden, die ihrerseits wiederum durch einen Second-order-Faktor bedingt werden würden. Um ein Higher-order-Faktormodell testen zu können, würde aber mehr als nur ein Item zur Messung des Faktors *Fixed Ability* benötigt. Aus Gründen der konzeptionellen Klarheit sollten die beiden Konstrukte aber getrennt werden. Dies würde bedeuten, dass im vorliegenden Messmodell das Item 02 zu entfernen ist, so dass sich alle Items wie bei Moschner und Gruber (im Druck) nur noch auf eine latente Variable beziehen. Die Höhe der Faktorladung des Items 02 (vgl. Tabelle 83) sprechen dafür, dass dieses Item nur marginal von diesem Faktor beeinflusst wird. Da das Messmodell inklusive Item 02 allerdings sowohl in der Analyse-, als auch der Validierungsstichprobe gute bis akzeptable Fitindizes aufwies, wurde auf einen Ausschluss dieses Items verzichtet

Tabelle 85: Fitindizes des Faktors Lernen lernen

Modell	N	χ^2 (df)	p	Bollen-Stine- p	χ^2/df	CFI	TLI	RMSEA	CI (90 %)	p_{close}	SRMR
LL	291	34.33 (9)	***	.01	3.82	.92	.86	.099	.065;.135	.01	.0507
LL final	291	8.95 (5)	.11	.22	1.79	.98	.97	.052	.000;.106	.40	.0323
LL S1	207	5.41	.37	.54	1.08	1.00	1.00	.020	.000;.100	.63	.0312
LL S2	221	9.20	.10	.22	1.84	1.00	.95	.062	.000;.124	.32	.0368
LL S3	217	8.89	.11	.12	1.78	.98	.95	.060	.0000;.123	.33	.0381
LL S4	212	7.47	.19	.28	1.49	.99	.97	.048	.0000;.114	.44	.0329
LL S5	221	6.81	.24	.37	1.36	.99	.98	.041	.0000;.108	.51	.0295
LL S6	215	8.80	.12	.25	1.76	.98	.96	.060	.0000;.123	.34	.0357
LL S7	217	9.76	.08	.17	1.95	.98	.95	.066	.0000;.128	.28	.0343
LL S8	221	6.06	.30	.46	1.21	.99	.98	.031	.0000;.103	.58	.0292
LL S9	219	13.25	.02	.05	2.65	.96	.92	.087	.031;.145	.12	.0452
LL S10	214	6.97	.22	.40	1.39	.99	.98	.043	.000;.111	.49	.0333
LL final	291	8.95 (5)	.11	.22	1.79	.98	.97	.052	.000;.106	.40	.0323
LL tau	291	134.80 (9)	***	.01	14.98	.45	.39	.220	.188;.253	***	.2034
LL paral- lel	291	215.55 (13)	***	.01	16.58	.12	.32	.232	.205;.260	***	.2179
LL Val.	126	8.49 (5)	.13	.22	1.70	.97	.93	.075	.000;.159	.26	.0529

Anmerkung: ***: $p < .001$

5.1.4 Beurteilung des revidierten Gesamtmodells

Nachdem nun die Modellpassung, faktorielle Validität sowie die psychometrischen Eigenschaften jedes Faktors des EPI-Fragebogens separat beurteilt wurden, soll das revidierte Gesamtmodell einer erneuten Prüfung auf Modellpassung unterzogen werden. Da das ursprüngliche Gesamtmodell keine ausreichende Modellpassung erzielte, wurde das Modell in jeweils einfaktorielle CFA-Modelle aufgeteilt. Nachdem jeder dieser Faktoren einzeln auf Modellpassung hin untersucht wurde und auftretende Fehlspezifikationen beseitigt wurden, soll nun das Ergebnis dieser einfaktoriellen CFAs wieder zu einem Gesamtmodell zusammengesetzt werden. Dabei werden die Items weggelassen, die sich in den einfaktoriellen CFA-Modellen als unpassend erwiesen haben. Ebenfalls ausgeschlossen wurden die Faktoren *Geschlechtsspezifische Wissenszugänge* und *Kulturspezifische Wissenszugänge*, da diese Faktoren interindividuelle Unterschiede in den Überzeugungen darstellen, dass sich der Umgang mit Wissen zwischen den Geschlechtern bzw. den Kulturen unterscheidet. Dies ist aber nicht, wie zuvor schon ausgeführt, mit Geschlechts- oder Kulturunterschieden in epistemologischen Überzeugungen, die eigentlich von Interesse wären, gleichzusetzen.

Das aus diesem Prozess resultierende Gesamtmodell umfasst sechs Faktoren, die Zuordnung der Items ist wie folgt:

- Faktor *Reflexive Natur des Wissens*: 07, 28, 48, 51,
- Faktor *Absolutheit des Wissens*: 15, 01, 29, 09, 17,
- Faktor *Sicherheit/Objektivität des Wissens*: 45, 13, 34, 44,
- Faktor *Soziale Komponente des Wissens*: 20, 22, 25, 52,
- Faktor *Value of Knowledge*: 12, 24, 38, 53,
- Faktor *Lernen lernen*: 02, 06, 31, 33, 47.

Um den Fit des Gesamtmodells zu testen, wurde ein 6-faktorielles CFA-Modell aufgestellt. Die Faktoren dürfen frei korrelieren, Skalierer sind jeweils die ersten in der obigen Zusammenstellung genannten Items. Korrelierende Fehlervariablen wurden nicht zugelassen, ebenso keine Doppelladungen eines Items auf einem anderen Faktor. Keine der Variablen des Modells lag bezüglich Schiefe und Wölbung über den Grenzen von West et al. (1995). Allerdings deutet ein signifikanter Mardia-Test darauf hin, dass die Variablen nicht multivariat normalverteilt sind ($M = 108.38$, $CR = 24.23$). Die Fitindizes dieses Modell sind $\chi^2 = 455.50$ ($df = 284$), $p < .01$, Bollen-Stine- $p = .01$, $\chi^2/df = 1.60$, CFI = .85, TLI = .84, SRMR = .068 und RMSEA = .046 (CI = [.038;.053], $p_{close} = .82$).

Diese Indizes lassen auf einen nicht akzeptablen Modellfit schließen. Allerdings ist zu bedenken, dass durch die Verletzung der multivariaten Normalverteilung sowie die relativ große Stichprobe selbst geringe Fehlspezifikationen zu signifikanten χ^2 -Werten führen können. Sowohl der RMSEA, dessen Konfidenzintervall relativ eng um den geschätzten Wert liegt, als auch der SRMR lässt vermuten, dass das grundlegende Modell nicht komplett von der empirischen Struktur der Daten abweicht.

Problematisch kann in dem aufgestellten Modell die Nichtbeachtung möglicher Doppelladungen von Items auf anderen Faktoren werden (Asparouhov & Muthén 2009). Da für das Gesamtmodell keine Doppelladungen zugelassen wurden, stellt sich die Frage, ob sich die Modellpassung bei ihrer Zulassung verbessern würde. Allerdings muss entschieden werden, welche Doppelladungen in das Modell aufgenommen werden sollen. Hierzu können die Modifikationsindizes zu Hilfe genommen werden. Grundsätzlich werden nur die Modifikationsindizes betrachtet, die einen Pfad von einer latenten Variablen zu einer manifesten Variablen vorschlagen.⁵⁰ Dieser soll groß genug sein, um den χ^2 -Wert des Modells signifikant zu verbessern und durch die Respezifikation entstehende Parameteränderung soll nach Möglichkeit statistisch bedeutsam sein. Hinweise auf eine signifikante Änderung des χ^2 -Wertes geben Modifikationsindizes, die größer 3.84 sind (dies entspricht einem auf dem 5%-Niveau signifikanten χ^2 bei $df=1$). Wesentlich schwerer zu beantworten ist die Frage, wie groß der im Rahmen der Modifikationsindizes geschätzte Parameterwert sein muss, um statistisch bedeutsam zu sein. Gemäß der Klassifikation von Cohen (1988) hat eine Korrelation dann eine mittlere Effektstärke, wenn sie den Wert .30 überschreitet. Da die Faktorladungen partielle Regressionsgewichte darstellen, wird für diese gefordert, dass mindestens eine mittlere Effektstärke vorliegt. Das ist gleichbedeutend mit der Forderung, dass die im Rahmen eines Modifikationsindex berechnete Änderung der Faktorladung größer als .30 ist, um für eine zu modellierende Doppelladung in Betracht zu kommen. Das Kriterium einer mittleren Effektstärke des partiellen Regressionsgewichts soll überdies sicherstellen, dass der modellierte Pfad inhaltlich sinnvoll zu interpretieren ist. Die relevanten Modifikationsindizes sind in Tabelle 86 zu finden. Diejenigen vorgeschlagenen Pfade, welche eine Bedingung des zuvor festgelegten Kriteriums nicht erfüllen, sind hervorgehoben.

⁵⁰ Die Aufnahme von Pfaden, die von einer manifesten Variablen zu einer anderen manifesten Variablen führen bzw. vorgeschlagene Pfade von Fehlervariablen zu latenten oder manifesten Variablen machen aus theoretischer Sicht keinen Sinn und werden deswegen nicht weiter beachtet. Diese theoretisch nicht sinnvollen Modifikationsindizes werden nur standardmäßig berechnet, da es dem Algorithmus aus naheliegenden Gründen nicht möglich ist, zwischen sinnvollen und sinnlosen Modifikationen zu unterscheiden.

Tabelle 86: Modifikationsindizes für das revidierte Gesamtmodell der EPI-Faktoren

Vorgeschlagene Pfade			Modifikationsindex	erwartete Parameteränderung
Item 02	-	VK	11.39	.82
Item 02	-	AW1	17.26	.52
Item 53	-	LL	8.02	-.95
Item 53	-	S/OW	10.97	.54
Item 53	-	AW1	12.17	.47
Item 53	-	RN	7.45	.45
Item 38	-	S/OW	6.36	-.40
Item 38	-	AW1	6.35	-.34
Item 12	-	LL	11.92	-.98
Item 12	-	SK	4.73	-.44
Item 12	-	S/OW	5.90	.33
Item 12	-	RN	10.23	.45
Item 52	-	AW1	4.80	.20
Item 20	-	LL	11.48	-1.11
Item 20	-	S/OW	8.65	.46
Item 20	-	AW1	4.06	.26
Item 20	-	RN	4.09	.32
Item 44	-	RN	4.14	-.39
Item 17	-	RN	7.54	.46
Item 48	-	S/OW	5.30	.21
Item 48	-	AW1	4.87	.17
Item 28	-	SK	5.75	-.33

Werden die aufgeführten Doppelladungen in das Modell aufgenommen, ergeben sich folgende Fitindizes: $\chi^2 = 341.04$ ($df = 266$), $p = .001$, Bollen-Stine- $p = .24$, $\chi^2/df = 1.28$, CFI = .94, TLI = .92, SRMR = .0445 sowie RMSEA = .031, CI = [.020;.041], $p_{close} = 1.00$. Die Faktoren des Modells und die Doppelladungen sind in Tabelle 87 dargestellt. Durch die Aufnahme der Doppelladung wird die Modellpassung signifikant verbessert, ein χ^2 -Differenztest ergibt $\Delta\chi^2 = 114.459$, $\Delta df = 18$, $p < .005$.

Die Fitindizes sind im Bereich guter bis akzeptabler Modellpassung. Der herkömmliche p-Wert liegt zwar unter der Grenze für eine akzeptable Modellpassung, allerdings ist auch hier die Sensibilität des χ^2 -Tests für geringe Fehlspezifikationen des Modells bei nicht normalverteilten Daten und die Abhängigkeit von der Stichprobengröße zu beachten. Der nach der Bollen-Stine-Bootstrap-Prozedur korrigierte p-Wert liegt im Bereich guter Modellpassung. Der CFI und TLI sind die beiden Indizes, die nur im Bereich akzeptabler Modellpassung liegen. Dies lässt drauf schließen, dass das Modell größtenteils zu der empirischen Struktur der Daten passt, wobei allerdings noch immer kleinere Fehlspezifikationen vorliegen.

Ergebnisse zum EPI

Tabelle 87: Faktorladung des Gesamtmodells

Faktor - Item Hauptladungen			partielles Regressionsgewicht		SE	CR	p
			nicht standardisiert	standardisiert			
Item 07	-	RN	1.00	.69			
Item 28	-	RN	.78	.49	.12	6.35	***
Item 48	-	RN	.79	.53	.11	7.41	***
Item 51	-	RN	1.06	.72	.12	9.11	***
Item 15	-	AW	1.00	.53			
Item 01	-	AW	1.18	.65	.17	6.80	***
Item 29	-	AW	1.14	.64	.17	6.73	***
Item 09	-	AW	.96	.53	.16	6.14	***
Item 17	-	AW	.76	.38	.16	4.82	***
Item 45	-	S/OW	1.00	.51			
Item 13	-	S/OW	1.42	.62	.27	5.30	***
Item 34	-	S/OW	.59	.32	.16	3.74	***
Item 44	-	S/OW	1.39	.54	.29	4.84	***
Item 20	-	SK	1.00	.44			
Item 22	-	SK	-1.16	-.38	.31	-3.69	***
Item 25	-	SK	-1.23	-.59	.29	-4.29	***
Item 52	-	SK	.93	.55	.22	4.26	***
Item 12	-	VK	1.00	.38			
Item 24	-	VK	2.33	.65	.58	4.05	***
Item 38	-	VK	2.71	.78	.70	3.85	***
Item 53	-	VK	1.42	.43	.39	3.60	***
Item 02	-	LL	1.00	.24			
Item 06	-	LL	-3.37	-.70	.95	-3.54	***
Item 31	-	LL	-2.25	-.49	.67	-3.36	***
Item 33	-	LL	-2.47	-.70	.70	-3.54	***
Item 47	-	LL	-2.25	-.62	.65	-3.49	***
Doppelladungen							
Item 02	-	VK	.43	.15	.23	1.90	.06
Item 02	-	AW	.39	.22	.14	2.84	.01
Item 53	-	LL	-.20	-.04	.40	-.49	.62
Item 53	-	S/OW	.16	.07	.21	.77	.44
Item 53	-	AW1	.37	.18	.18	2.10	.04
Item 53	-	RN	.45	.18	.22	2.04	.04
Item 38	-	SO/W	-.02	-.01	.23	-.11	.92
Item 38	-	AW	-.36	-.17	.22	-1.66	.10
Item 12	-	LL	-.50	-.13	.37	-1.36	.18
Item 12	-	SK	-.27	-.14	.21	-1.26	.21
Item 12	-	SO/W	.13	.07	.15	.86	.39
Item 12	-	RN	.15	.08	.19	.79	.43
Item 20	-	LL	-.85	-.19	.45	-1.87	.06
Item 20	-	S/OW	.36	.17	.18	2.03	.04
Item 20	-	RN	.10	.04	.23	.42	.68
Item 44	-	RN	-.50	-.18	.22	-2.32	.02
Item 17	-	RN	.46	.19	.16	2.83	.01
Item 28	-	SK	-.39	-.26	.14	-2.90	***

Anmerkung: ***: $p < .001$

Tabelle 88: Geschätzte Korrelation der Faktoren im Gesamtmodell

	RN	AW	S/OW	SK	VK
AW	-.12*	-			
S/OW	.27**	.32**	-		
SK	-.41**	.26**	-.17**	-	
VK	-.20**	.32**	-.05	.47**	-
LL	-.51**	-.12*	-.28**	.25**	.03

Anmerkung: *: $p < .05$, **: $p < .001$

Die Faktorladungen zeigen, dass bis auf Item 02 jede der manifesten Variablen aus den ein-faktoriellen CFA-Modellen auch im Gesamtmodell substantziell auf den zugeordneten Faktor lädt. Ebenso sind alle Faktorladungen signifikant. Die in das Modell aufgenommenen Kreuz-ladungen sind alle nicht substantziell, obwohl einige signifikant sind.

Eine weitere Möglichkeit der Evaluierung des Gesamtmodells besteht darin, dieses mithilfe eines E/CFA-Modell darzustellen. Zu diesem Zweck wurde eine EFA mit Maximum-Likelihood-Extraktion von sechs Faktoren mit anschließender Promax-Rotation ($\kappa = 4$) durchgeführt. Das KMO-Maß betrug .73. Die Mustermatrix ist in Tabelle 89 und die Korrela-tionsmatrix der Faktoren ist in Tabelle 90 aufgeführt.

Tabelle 89: Mustermatrix der ML-EFA

	Faktor					
	1	2	3	4	5	6
Item 07 -RN	-.01	-.06	.71	.09	-.03	.04
Item 28 - RN	.04	-.04	.54	-.08	.09	.04
Item 48 - RN	.19	.05	.38	-.01	.04	.07
Item 51 - RN	.03	.03	.77	-.03	-.05	-.12
Item 15 - AW	.18	.53	-.20	-.08	-.04	-.04
Item 01 - AW	.13	.58	-.12	.05	.06	.05
Item 29 - AW	-.11	.73	.11	-.01	-.09	-.11
Item 09 - AW	-.05	.54	-.01	-.04	.09	.06
Item 17 - AW	-.09	.36	.21	.02	.03	.08
Item 45 -S/OW	-.02	-.05	.09	.06	.00	.49
Item 13 - S/OW	.02	.04	.01	-.07	-.02	.57
Item 34 - S/OW	-.03	.11	.07	.00	.16	.22
Item 44 - S/OW	-.04	-.02	-.13	-.03	-.03	.61
Item 20 - SK	.10	.04	.00	.20	-.20	.16
Item 22 - SK	-.03	.13	.08	-.21	.24	-.02
Item 25 - SK	.02	.01	-.03	.05	1.02	.01
Item 52 - SK	-.09	.20	-.11	.13	-.25	.04
Item 12 - VK	.14	.03	.10	.28	.01	.03
Item 24 - VK	.01	.05	-.09	.59	.02	-.06
Item 38 - VK	-.05	-.13	-.01	.81	.02	-.05
Item 53 - VK	.07	.17	.13	.42	-.05	.06

Ergebnisse zum EPI

	Faktor					
	1	2	3	4	5	6
Item 02 - LL	-.26	.27	-.01	.15	.12	-.03
Item 06 - LL	.69	.02	.03	.04	-.03	-.07
Item 31 - LL	.43	.01	.09	.02	-.04	.09
Item 33 - LL	.70	.05	.05	-.01	-.01	-.09
Item 47 - LL	.66	-.07	-.06	.04	.11	.04

Tabelle 90: Korrelationsmatrix der Faktoren

Faktor	RN	AW	S/OW	SK	VK
AW	.11	-			
S/OW	.45**	-.08	-		
SK	.00	.29**	-.14*	-	
VK	.15*	-.12*	.31**	-.24**	-
LL	.29**	.34**	.20**	.05	.11

Anmerkung: *: $p < .05$, **: $p < .001$

Die Mustermatrix demonstriert, dass für die Faktoren RN, AW, S/OW, VK und LL eine klare Faktorenstruktur existiert. Auch bei der EFA weist Item 02 des Faktors LL eine nur geringe Hauptladung und eine noch größere Nebenladung auf. Der Faktor SK hingegen weist keine homogene Ladungsstruktur auf. Auffällig ist auch die Ladung > 1 von Item 25 auf dem fünften Faktor. Solche Ladungen können bei Promax-Rotationen vorkommen und verschwinden in der Regel bei Wahl eines anderen κ , deuten aber auf eventuelle Probleme mit der betreffenden Variablen hin (Tataryn, Wood & Gorsuch, 1999).

Die mittels EFA ermittelten Faktorladungen wurden in ein E/CFA-Modell eingesetzt, wobei die Mustermatrix der EFA die Λ_X -Matrix und die Korrelationsmatrix der Faktoren die Φ -Matrix der CFA darstellt. Für dieses Modell wurde die Modellpassung berechnet. Das Modell zeigte eine gute Passung ($\chi^2 = 341.57$ ($df = 325$), $p = .25$, Bollen-Stine- $p = .80$, $\chi^2/df = 1.05$, CFI = .99, TLI = .99, SRMR = .0509, RMSEA = .013, CI = [0;.026], $p_{close} = 1.00$), lediglich der SRMR lag nur ganz knapp außerhalb der Grenzen für eine gute Modellpassung.

Da bei einer EFA die Ergebnisse immer durch die Wahl der Extraktions- und Rotationsmethode bestimmt werden, wurde die EFA unter Benutzung der Hauptachsenanalyse mit Promax-Rotation ($\kappa = 4$) wiederholt. Die Mustermatrix und die Korrelationsmatrix der Faktoren sind in den folgenden Tabellen aufgeführt.

Ergebnisse zum EPI

Tabelle 91: Mustermatrix der PAF-EFA

	Faktor					
	1	2	3	4	5	6
Item 07 - RN	-.02	-.06	.73	.07	-.04	.03
Item 28 - RN	.05	-.02	.50	-.06	.19	.04
Item 48 - RN	.21	.06	.35	-.03	.05	.06
Item 51 - RN	.03	.03	.75	-.05	-.02	-.13
Item 15 - AW	.19	.54	-.20	-.05	.00	-.06
Item 01 - AW	.14	.58	-.09	.04	-.03	.06
Item 29 - AW	-.11	.71	.10	.02	-.01	-.10
Item 09 - AW	-.05	.54	.02	-.07	.04	.08
Item 17 - AW	-.09	.33	.21	.01	-.01	.10
Item 45 - S/OW	-.04	-.06	.06	.07	-.02	.52
Item 13 - S/OW	.03	.04	-.04	-.02	.04	.57
Item 34 - S/OW	-.03	.10	.10	-.01	.06	.24
Item 44 - S/OW	-.03	-.01	-.14	-.07	-.06	.59
Item 20 - SK	.10	.03	.06	.04	-.41	.17
Item 22 - SK	-.02	.14	.02	-.05	.44	.00
Item 25 - SK	.03	.00	.05	.08	.63	.10
Item 52 - SK	-.09	.21	-.06	-.02	-.42	.01
Item 12 - VK	.14	.02	.05	.38	.12	.03
Item 24 - VK	.02	.02	-.13	.72	.08	-.04
Item 38 - VK	-.06	-.12	.03	.65	-.18	-.03
Item 53 - VK	.07	.15	.15	.36	-.17	.07
Item 02 - LL	-.25	.28	-.05	.21	.20	-.03
Item 06 - LL	.68	.03	.03	.06	-.01	-.07
Item 31 - LL	.44	.02	.08	.00	-.06	.06
Item 33 - LL	.69	.05	.06	-.04	-.05	-.08
Item 47 - LL	.67	-.08	-.08	.08	.10	.05

Tabelle 92: Korrelationsmatrix der Faktoren

Faktor	RN	AW	S/OW	SK	VK
AW	.09	-			
S/OW	.46**	-.08	-		
SK	.01	.28**	-.08	-	
VK	.15*	-.22**	.31**	-.34**	-
LL	.31**	.32**	.26**	.04	.09

Anmerkung: *: $p < .05$, **: $p < .001$

Auch diese Mustermatrix lässt erkennen, dass für die Faktoren RN, AW, S/OW, VK und LL eine klare Faktorenstruktur existiert. Anders als bei der ML-Extraktion ergibt sich aber auch für den Faktor SK eine klare Zuordnung der Items zu dem zugehörigen Faktor. Ebenfalls wie bei der ML-Extraktion weist Item 02 des Faktors LL eine nur geringe Hauptladung und eine noch größere Nebenladung auf.

Analog zum obigen Vorgehen wurde die Mustermatrix in die Λ_X -Matrix und die Korrelationsmatrix der Faktoren in die Φ -Matrix der CFA eingesetzt. Das E/CFA-Modell zeigte eine gute Passung ($\chi^2 = 344.62$ ($df = 325$), $p = .22$, Bollen-Stine- $p = .78$, $\chi^2/df = 1.06$, CFI = .99, TLI = .99, SRMR = .0501, RMSEA = .013, CI = [0;.026], $p_{close} = 1.00$).

In den beiden durchgeführten EFAs lassen sich überdies mehrere der während der Modellmodifikation in das Modell aufgenommene Doppelladungen finden, wobei aber auch hier diese Ladungen gering ausfallen (im Folgenden werden nur Doppelladungen größer als .20 beachtet). Bei der Maximum-Likelihood-Faktorenanalyse findet sich die Doppelladung von Item 17 auf dem Faktor *Reflexive Natur des Wissens* und von Item 02 auf dem Faktor *Absolutes Wissen*. Bei der Hauptachsenanalyse finden sich die Doppelladung des Items 48 auf dem Faktor *Lernen lernen*, die Doppelladungen der Items 15 und 17 auf dem Faktor *Reflexive Natur des Wissens* sowie die Doppelladungen des Items 02 auf den Faktoren *Absolutes Wissen* und *Value of Knowledge*. Auch das uneinheitliche Verhalten des Faktors *Soziale Komponente des Wissens* unter den beiden Extraktionsverfahren muss geklärt werden, es konnte durch diese beiden Faktorenanalysen nicht bestimmt werden, ob dieser Faktor eine eigenständige latente Variable ist oder ob die ihm zugeordneten Items durch andere Faktoren beeinflusst werden. Der Befund stützt die schon im Zusammenhang mit dem einfaktoriellen Messmodell dieses Faktors geäußerte Vermutung, dass der Faktor durch ein *Higher-order*- oder hierarchisches Faktormodell besser modelliert werden kann. Item 02 des Faktors *Lernen lernen* stellte sich schon in dem einfaktoriellen CFA-Modell als problematisch heraus, sodass die Frage besteht, ob das Item aufgrund seines Inhalts nicht aus dem Modell zu entfernen ist.

5.2 Diskussion der Analysen des EPI

Im Folgenden werden einige Besonderheiten, die sich bei Analyse der EPI-Skalen ergeben haben, diskutiert. Das Hauptaugenmerk liegt dabei auf dem Vergleich mit dem FEE-Fragebogen von Moschner und Gruber (im Druck) sowie den methodischen Besonderheiten.

Die erste Forschungsfrage zielte auf die Passung des Gesamtmodells der EPI-Faktoren. Das Gesamtmodell zeigte zu Beginn keine ausreichende Modellpassung. In methodischer Hinsicht kommt die Art der bei der Skalenkonstruktion benutzten faktorenanalytischen Auswertung als mögliche Erklärung für die mangelnde Modellpassung in Betracht (vgl. Kapitel 1.3.3 für die Darstellung der Entwicklung des EPI-Fragebogens). Moschner und Gruber (2005) geben an, zur Ermittlung der dimensional Struktur eine Hauptkomponentenanalyse durchgeführt zu haben (bei der Weiterentwicklung bleibt unklar, ob eine Hauptkomponentenanalyse oder eine Hauptachsenanalyse durchgeführt wurde, Moschner und Gruber (2005a, 2005b) sprechen nur von Faktorenanalyse). Hauptkomponentenanalysen stellen allerdings ein reines Instrument zur Dimensionsreduktion dar (Backhaus et al., 2003). Der Hauptunterschied zur Hauptachsenanalyse ist, dass bei der Hauptkomponentenanalyse keine Messfehler berücksichtigt werden (Moosbrugger & Hartig, 2002, Russell, 2002). Legt man das *Common-factor*-Modell von Thurstone (1947) zugrunde, so würde dort der Fehlerterm wegfallen. Damit entfällt aber auch die theoretische Annahme, dass die Varianz einer manifesten Variablen in die Varianz, die durch einen oder mehrere gemeinsame Faktoren verursacht wird, und in die Residualvarianz aufteilen lässt. Somit wird bei der Berechnung einer Hauptkomponentenanalyse der Fehlerterm nicht berücksichtigt, dies führt aber zu einer Vermischung der Varianz, die durch die Varianz der gemeinsamen Faktoren bestimmt wird, und der Residualvarianz. Hierdurch kann es zu Verzerrungen der Faktorladungen kommen, häufig werden die Komponentenladungen und die Korrelationen zwischen den Komponenten überschätzt (Widaman, 1993, Fabrigar, Wegener, MacCallum & Strahan, 1999, Preacher & MacCallum, 2003). Die Verwendung der Hauptkomponentenanalyse kann also zu Ergebnissen geführt haben, die sich erheblich von denen unterscheiden, die bei der Verwendung einer explorativen Faktorenanalyse mit Berücksichtigung der Residualvariablen ergeben hätte. Allgemein gilt überdies auch, dass sich Faktorstrukturen, die mittels explorativer Faktorenanalysen ermittelt wurden, schlecht mithilfe konfirmatorischer Faktorenanalysen bestätigen lassen (Van Prooijen & Van der Kloot, 2001). Zudem werden bei einer CFA üblicherweise nur Ladungen auf jeweils einem Faktor berücksichtigt, während bei einer EFA jedes Item auf jeweils allen Faktoren lädt. Das CFA-Modell bildet somit nur eine Einfachstruktur ab, die auch als *independent cluster model* bezeichnet

wird. Komplexere Ladungsmuster wurden somit im *independent cluster modell* nicht berücksichtigt und in der Regel zeigen diese daher eine mangelnde Modellpassung (Asparouhov & Muthén, 2009). Aufgrund der schlechten Passung des initialen Gesamtmodells wurde anschließend die Modellpassung jedes Faktors einzeln geprüft.

Der erste Faktor, der geprüft wurde, war *Reflexive Natur des Wissens*, dem fünf Items zugeordnet wurden. Der Faktor erzielte zu Beginn eine gute bis akzeptable Modellpassung, die aber durch den Ausschluss von Item 27 deutlich verbessert werden konnte. In FEE von Moschner und Gruber (im Druck) sind alle Items der Skala *Reflexive Natur des Wissens* enthalten. Ob der Ausschluss des Items 27 somit gerechtfertigt war oder ob eine Überanpassung des Modells an die vorliegenden Daten der Fall war, lässt sich nicht mit Sicherheit feststellen. Dies kann aber nur in weiterführenden Untersuchungen mit neuen Stichproben erfolgen. Die faktorielle Validität des Faktors *Reflexive Natur des Wissens* lässt sich an dieser Stelle somit nicht abschließend beurteilen. Für die weiterführenden Analysen, insbesondere dem später untersuchten Zusammenhang von epistemologischen Überzeugungen mit Lernstrategien, wird das einfaktorielle Messmodell ohne Item 27 benutzt. Das vorliegende einfaktorielle CFA-Modell sowie die Ergebnisse der Regression der EPI-Dimensionen auf Lernstrategien haben also nur vorläufigen Charakter und bedürfen weiterer Bestätigung. Bei dem Faktor *Reflexive Natur des Wissens* verdeutlichen die Ergebnisse der Validierung, wie wichtig die Aspekte der Itemgestaltung sind. Als problematisch hat sich erwiesen, dass Item 27 und 28 inhaltlich ähnlich sind. Besonders problematisch hierbei war, dass wegen der fehlenden Modellpassung der Validierungsstichprobe keine eindeutige Festlegung getroffen werden konnte, welches der beiden Items besser zur Erfassung des Faktors *Reflexive Natur des Wissens* geeignet ist. Gezeigt werden konnte aber sicherlich, dass es möglich ist, den Faktor mit den vorliegenden Items zu erfassen, auch wenn die Ergebnisse vorerst nur im Kontext der untersuchten Stichprobe betrachtet werden können.

Der nächste Faktor war *Absolutes Wissen*. Dieser konnte mithilfe einer EFA in die zwei Faktoren *Absolutheit des Wissens* und *Sicherheit/Objektivität des Wissens* untergeteilt werden, wobei sich aber die Kreuzvalidierung des zweiten Faktors als problematisch erwies. Der zweite Faktor *Sicherheit/Objektivität des Wissens* ist vorläufig nur im Rahmen der Analysestichprobe gültig. Analog dem Faktor *Reflexive Natur des Wissens* wird der Faktor *Sicherheit/Objektivität des Wissens* für die Regression epistemologischer Überzeugungen auf Lern-

strategien genutzt, ebenfalls unter der Einschränkung, dass die Ergebnisse nur für die vorliegende Stichprobe gelten.

Im FEE von Moschner und Gruber (im Druck) wurde der gemeinsame Faktor *Absolutes Wissen* ebenfalls in zwei Faktoren aufgeteilt. Die beiden Faktoren im FEE sind *Umgang mit Autoritäten* und *Sicherheit von Wissen*, diese entsprechen weitgehend den beiden Faktoren *Autoritäres Wissen* und *Sicherheit/Objektivität des Wissens*. Die Itemzuordnungen der Faktoren werden in Tabelle 93 einander gegenübergestellt.

Tabelle 93: Vergleich der EPI- und FEE-Faktoren

	EPI		FEE	
	Absolutes Wissen	Sicherheit/Objektivität des Wissens	Umgang mit Autoritäten	Sicherheit des Wissens
	01	13	04	13
	04	16	09	17
	09	34	15	34
Item Nr.	15	37	16	37
	17	44	29	44
	29	45	42	45
				49

Die EPI- und FEE-Faktoren weisen somit inhaltlich eine große Übereinstimmung auf. Die Teilung des ursprünglichen Faktors *Absolutes Wissen* in die beiden Faktoren *Autoritäres Wissen* und *Sicherheit/Objektivität des Wissens* kann somit durch die Ergebnisse von Moschner und Gruber als konzeptionell repliziert angesehen werden. Allerdings sind die Befunde der Kreuzvalidierung zu bedenken, sodass eine Übertragung der Ergebnisse auf andere Stichproben weiter untersucht werden muss.

Der nächste zu prüfende Faktor war *Soziale Komponente des Wissens*. Das initiale Messmodell zeigt keinen ausreichenden Modellfit und es wurden in zwei Schritten die Items 40 und 41 entfernt. Das resultierende Messmodell besaß eine hervorragende Modellpassung. Der Faktor findet sich auch im FEE, allerdings mit anderen Items. Von Moschner und Gruber (2007) wurde Item 22 entfernt, während die Items 40 und 41 in der Skala belassen wurden.

Bei dem Faktor *Value of Knowledge* zeigte das initiale Messmodell ebenfalls keine ausreichend hohe Modellpassung. Der Ausschluss von Item 21 führte zu einem guten Modellfit. Das Modell wies auch in der Validierungstichprobe eine gute Modellpassung auf, diese war sogar noch besser als in der Analysetichprobe. Von daher ist also nicht zu erwarten, dass das

revidierte Modell an die Daten der Analysestichproben angepasst wurde. Bei Moschner und Gruber (im Druck) enthält das Messmodell zusätzlich Item 21 und ist somit identisch mit dem ursprünglichen Messmodell dieses EPI-Faktors.

Der nächste Faktor ist *Geschlechtsspezifische Wissenszugänge*. Hier zeigte das einfaktorielle Messmodell eine sehr schlechte Passung. Die Items, die diesen Faktor erfassen sollen, weisen fast alle Formulierungen auf, die sich auf die Gleichheit bzw. Verschiedenheit der Geschlechter beziehen. Ähnlichkeiten in der Formulierung können zu Methodenfaktoren führen. Diese können bei einer explorativen Analyse zu zweidimensionalen Strukturen führen. Z. B. wurde anhand der *Self-Esteem*-Skala von Rosenberg (1965) nachgewiesen, dass die beiden Faktoren positives und negatives Self-Esteem nichts anderes als Methodenartefakte darstellen (Marsh, 1996, Tomás & Oliver, 1999). Ähnliches wurde auch für den *Penn State Worry Questionnaire* (Meyer, Miller, Metzger & Borkovex, 1990) nachgewiesen. Dieser enthält zwei Faktoren, deren Items in symptomatischer und nicht symptomatischer Richtung formuliert sind. Auch diese Faktorenlösung erwies sich als ein Artefakt der Itemformulierung (Brown, 2006, Hazlett-Stevens, Ulman & Craske, 2004). In beiden Beispielen handelt es sich jeweils um eindimensionale Faktoren, denen entsprechende Methodenfaktoren zugeordnet sind. Der Methodenfaktor gibt in diesem Fall Einflüsse der gleichartigen Itemformulierung als latente Variable auf die manifesten Variablen an. Es handelt sich dabei um eine Quelle systematischer Varianz, die neben der eigentlich interessierenden latenten Variablen die manifesten Variablen beeinflusst.

Bei der Skalenkonstruktion im FEE von Moschner und Gruber (im Druck) besteht dieser Faktor aus den Items 23, 30, 36, 43 und 46. Dies sind diejenigen Items, die sich auf die Unterschiedlichkeit der Geschlechter beziehen. Lediglich Item 08, das sich ebenfalls auf die Unterschiedlichkeit von Männern und Frauen bezieht, ist nicht in deren Messmodell vorhanden. Überdies sind dies auch alle Items, die in der durchgeführten EFA zusammen auf dem ersten Faktor luden. Die Teilung der Items durch die Formulierung in solche, die sich auf die Unterschiedlichkeit bzw. Gleichheit von Männern und Frauen im Umgang mit Wissen beziehen, erweist sich als sehr ungünstig.

Viel bedeutsamer als diese Methodeneffekte sind allerdings inhaltliche Bedenken. Epistemologische Überzeugungen kennzeichnen den Umgang einer Person mit Wissen. In dem persönlichkeitspsychologischen Ansatz, der hier angewandt wird, geht es um interindividuelle Un-

terschiede in den epistemologischen Überzeugungen von Personen. Der Faktor *Geschlechtsspezifische Wissenszugänge* hingegen bezieht sich nicht auf die interindividuellen Unterschiede zwischen Personen im Umgang mit Wissen, sondern er bezieht sich auf interindividuelle Unterschiede im Denken über Geschlechtsunterschiede im Umgang mit Wissen. Der Faktor erfasst also nicht etwa vorhandene Geschlechtsunterschiede in den eigentlich interessierenden epistemologischen Überzeugungen. Interindividuelle Unterschiede in dieser Dimension beantworten also nicht die Frage, ob die epistemologischen Überzeugungen Geschlechtsunterschiede zeigen. Zudem müsste zu der Beantwortung dieser Frage zusätzlich zu einer entsprechenden Gestaltung der Items ein Untersuchungsdesign gewählt werden, welches zur Erfassung von Geschlechtsunterschieden geeignet ist. Somit hat der Faktor *Geschlechtsspezifische Wissenszugänge* keinen Bezug zu epistemologischen Überzeugungen. Ob und inwiefern das Denken über Geschlechtsunterschiede im Umgang mit Wissen im Zusammenhang mit epistemologischen Überzeugungen steht, bleibt an dieser Stelle unklar. Wegen des fehlenden Bezugs zu epistemologischen Überzeugungen und der methodischen Probleme wird der Faktor in der weiteren Auswertung nicht berücksichtigt.

Der Faktor *Kulturspezifische Wissenszugänge* wies beim initialen Messmodell auch eine ungenügende Modellpassung auf, sodass auch bei diesem Faktor eine Revision des Messmodells erforderlich war. Allerdings führte das revidierte Messmodell in der Validierungsstichprobe zu keiner ausreichenden Modellpassung. Bei diesem Faktor stellte sich aber das gleiche Problem wie bei dem Faktor *Geschlechtsspezifische Wissenszugänge*. Der Faktor erfasst keine kulturellen Unterschiede epistemologischer Überzeugungen, sondern lediglich interindividuelle Unterschiede im Denken über den Wissenserwerb in verschiedenen Kulturen. Interindividuelle Unterschiede in dieser Dimension beantworten also nicht die Frage, ob es kulturelle Unterschiede epistemologischer Überzeugungen gibt.

Um Geschlechtsunterschiede bzw. Kulturunterschiede in epistemologischen Überzeugungen zu erfassen, werden entsprechende Methoden benötigt. Zur Erfassung von Geschlechtsunterschieden bietet sich z. B. eine getrennte Erfassung epistemologischer Überzeugungen an. Dieser Ansatz wurde von Belenky et al. (1986) und Baxter Magolda (1992) im Rahmen der auf Perry (1968) zurückgehenden Stufenmodelle verfolgt und ist mit den für diesen Ansatz typischen Methoden der Interviewstudien durchgeführt. Belenky et al. haben sich dabei ausschließlich auf die Erfassung epistemologischer Überzeugungen bei Frauen konzentriert, was von Baxter Magolda (1992) als zu einseitig kritisiert wurde und mit einem Vergleich der

Entwicklungsstufen, die jeweils bei Männern und Frauen erhoben wurden, erweitert. In dem Paradigma der Erfassung epistemologischer Überzeugungen mithilfe von Fragebögen ist eine entsprechende Methodik des Vergleichs der erfassten Faktoren zwischen Männern und Frauen erforderlich. Im Rahmen von konfirmatorischen Faktorenanalysen können hierbei z. B. Multigruppenanalysen durchgeführt werden, wobei die Untersuchungsgruppen durch das Geschlecht definiert werden. Von Interesse ist in diesem Fall, ob sich die Mittelwerte der latenten Variablen zwischen den Geschlechtern unterscheiden. Um solch einen Gruppenvergleich durchführen zu können, sind allerdings einige Voraussetzungen zu beachten. So ist es erforderlich, dass die faktorielle Invarianz überprüft wird (vgl. Meredith, 1993, Vandenberg & Lance, 2000). Diese teilt sich in acht sukzessive Schritte, die sicherstellen, dass in den beiden Gruppen jeweils das gleiche Konstrukt erfasst wird, der Ablauf wird in Anlehnung an Brown (2006) in Tabelle 93 dargestellt.

Tabelle 94: Schritte bei der Prüfung faktorieller Invarianz

Schritt	Test	Ebene
1	CFA-Modell in allen Gruppen	
2	Gleichheit der Faktorenstruktur	
3	Gleichheit der Faktorladungen	Messinvarianz
4	Gleichheit der Indikatorkonstanten	
5	Gleichheit der Indikatorresiduen	
6	Gleichheit der Faktorvarianzen	
7	Gleichheit der Faktorkovarianzen	Populationsheterogenität
8	Gleichheit der Mittelwerte der latenten Variablen	

Dabei müssen die Tests sequenziell abgearbeitet sein, da der Test bei einem Schritt jeweils von den vorhergehenden Tests abhängt. So ist z. B. die Prüfung auf Gleichheit der Faktorenstrukturen in Schritt 2 davon abhängig, dass in beiden Gruppen das Messmodell eine ausreichende Passung aufweist.

Dabei wird in den Schritten 1 bis 5 die Messinvarianz getestet. Messinvarianz bedeutet, dass in allen betrachteten Gruppen das Messmodell eine akzeptable Modellpassung aufweist und dass die Faktorstruktur, die Faktorladungen, die Konstanten der Indikatoren sowie deren Residualvarianzen gleich sind. Sind alle diese Bedingungen erfüllt, ist sichergestellt, dass in allen Gruppen dasselbe Konstrukt erfasst wird. In den Schritten 5 bis 8 werden die Gleichheit der Faktorvarianzen und -kovarianzen sowie die eingangs erwähnten Mittelwerte der latenten Variablen auf Gleichheit geprüft. Dabei wird in diesen Schritten nicht die Äquivalenz des Messinstruments, sondern die Heterogenität der jeweiligen Faktoren in den einzelnen Grup-

pen untersucht. Die Gleichheit bzw. Unterschiedlichkeit der Mittelwerte der latenten Variablen gibt dann Auskunft über Geschlechtsunterschiede, wobei Gruppenvergleiche der latenten Mittelwerte sind nur dann sinnvoll sind, wenn die Indikatorkonstanten sowie die Faktorladungen in beiden Gruppen gleich sind (Brown, 2006). Ein besonderer Fall beim Vergleich der latenten Mittelwerte ist das *differential item functioning* (DIF). Hierbei weisen die Items unterschiedliche latente Mittelwerte in den untersuchten Gruppen auf (McDonald, 1999).

Das sequenzielle Prüfen der dargestellten Schritte mit den Gruppen Männer und Frauen gestattet es also zum einen, Aussagen über die Äquivalenz des Messinstruments (Messinvarianz) und zum anderen Aussagen über Unterschiede in der Faktorenstruktur zwischen den Gruppen zu treffen (Populationsheterogenität). Dieses Vorgehen erlaubt somit eine differenzierte Betrachtung von Geschlechtsunterschieden epistemologischer Überzeugungen.

Die vorgestellte Methodik ist aber nicht nur zur Untersuchung von Geschlechtsunterschieden, sondern auch von kulturellen Unterschieden epistemologischer Überzeugungen geeignet. Hierbei werden die Gruppen nicht durch die Geschlechter, sondern durch die zu untersuchenden Kulturen gebildet. Dabei ist es nicht notwendig, die Betrachtung auf nur zwei Gruppen (Kulturen) zu beschränken. Im Rahmen der Mehrgruppen-CFA können beliebig viele Gruppen gegenübergestellt werden. Bei der Untersuchung von kulturellen Unterschieden müssen aber auch kulturspezifische Besonderheiten berücksichtigt werden. Nach Pike (1954) unterscheidet man dabei die in den Kulturwissenschaften benutzte Typisierung etisch und emisch. Etische Konstrukte sind kulturübergreifend vorhanden, während emische Konstrukte nur für eine Kultur relevant sind und in anderen Kulturen keine besondere Bedeutung besitzen, für psychologische Konstrukte wurde diese Unterscheidung von Leung und Bond (1989) übernommen. Eine Mehrgruppen-CFA kann auch hierrüber Auskunft geben. Etische Konstrukte würde zu einer Messinvarianz in verschiedenen Kulturen führen, während emische Konstrukt nicht messinvariant wären.

Die Untersuchung von Geschlechts- und kulturellen Unterschieden epistemologischer Überzeugungen ist also mithilfe der beiden dargestellten Faktoren *Geschlechtsspezifische Wissenszugänge* und *Kulturspezifische Wissenszugänge* nicht möglich. Vielmehr sind hier, wie aufgezeigt wurde, andere methodische Herangehensweisen erforderlich. Auf diesem Grund wurden und werden die beiden Faktoren nicht weiter berücksichtigt.

Der letzte Faktor, der untersucht wurde, war *Lernen lernen*. Das ursprüngliche Messmodell musste auch hier in einem Schritt modifiziert werden, um eine ausreichende Modellpassung zu erzielen. Das Modell zeigte auch in der Validierungsstichprobe eine gute Modellpassung, sodass nicht von einer Überanpassung des Messmodells an die Daten der Analysestichprobe ausgegangen werden kann. Der Faktor bildete eine Mischung aus den beiden hypothetischen Konstrukten *Fixed Ability* und *Strategische Komponente des Lernens* (vgl. Kapitel 1.3.3). Das kritische Item hierbei war Item 02, das sich Begabung bezieht. Bei Moschner und Gruber (im Druck) hingegen finden sich nur Items, die sich auf den Erwerb entsprechender Lerntechniken, also die strategische Komponente des Lernens beziehen. Insofern ist an dieser Stelle die Frage zu stellen, ob es aus theoretischer Sicht nicht sinnvoll wäre, das entsprechende Item 02 zu entfernen.

Der Faktor *Lernen lernen* erlaubt die Erfassung individueller Unterschiede im Hinblick auf das Denken über den Erwerb von Lernmethoden. Allerdings muss auf eine Besonderheit des Faktors hingewiesen werden. Der Faktor selbst erfasst keine interindividuellen Differenzen beim Erwerb von Lernmethoden, sondern lediglich die Einstellung über die Möglichkeit des Erwerbs von Lernstrategien. Daher ist die Beziehung dieses Faktors zu den eigentlichen Lernstrategien von besonderem Interesse. Dieser Frage wird im Rahmen der hier untersuchten Fragestellungen im nächsten Kapitel nachgegangen.

Im Anschluss an die Prüfung der einzelnen Faktoren epistemologischer Überzeugungen wurden die modifizierten Faktoren erneut zu einem Gesamtmodell zusammengesetzt. Auch dieses Gesamtmodell wurde als *independent cluster model* konzipiert. Das revidierte Gesamtmodell zeigte ebenfalls eine mangelnde Modellpassung. Da die einzelnen Faktoren auf ihre Modellpassung untersucht wurden und die revidierten Messmodelle jeweils mindestens akzeptable Modellpassung aufwiesen, ist der Grund für die mangelnde Passung wahrscheinlich in der Einfachstruktur des Gesamtmodells zu sehen. Um durch das Fehlen von Kreuzladungen bedingte mögliche Fehlanpassungen des Gesamtmodells zu erkennen, wurden die Modifikationsindizes betrachtet. Diese Strategie findet in der zuvor beschriebenen Situation eines schlecht passenden CFA-Modells mit der Annahme einer Einfachstruktur häufig Anwendung und wurde von MacCallum et al. (1992) wegen der häufig resultierenden *specification searches* kritisiert. Aus diesem Grund wurde hier eine doppelte Strategie verfolgt. Zum einen wurden nur solche Modifikationsindizes betrachtet, die eine signifikante Verbesserung der Modellpassung im Sinne des χ^2 -Wertes nahelegen. Dies war das in der Literatur übliche Kriterium

eines Modifikationsindex größer als 3.84 (vgl. Bühner, 2006). Zum anderen sollten aus dieser eingeschränkten Menge aus Modifikationsindizes nur solche ausgewählt werden, die einen bedeutenden Parameterwert aufwiesen. Hier wurden Anleihen bei der Klassifikation von Cohen (1988) gemacht und nur solche Modifikationsindizes betrachtet, die einen geschätzten Parameterwert im Sinne eines mittleren Effekts aufwiesen. Auf diese Weise konnten 18 mögliche Modellmodifikationen identifiziert werden, die mögliche Nebenladungen nahelegten. Nachdem diese Nebenladungen in das Modell aufgenommen worden waren, zeigten sich gute bis akzeptable Fitindizes. Ein χ^2 -Differenztest weist auf eine signifikante Verbesserung der Modellpassung hin. Auf diese wurden 18 Nebenladungen in das Modell eingebracht, von denen aber nur sieben signifikant waren und keine der standardisierten Regressionskoeffizienten letztendlich die geforderte Mindestgröße von .30 erreichte.

Eine weitere Möglichkeit, die Modellpassung des Gesamtmodells zu evaluieren, besteht darin, ein E/CFA-Modell zu prüfen (Asparouhov & Muthén, 2009). Hierzu und um evtl. vorhandene Nebenladungen bestimmter Items abzuklären, wurden die Items mithilfe zweier explorativer Faktorenanalysen untersucht, wobei einmal eine Hauptachsenanalyse und einmal eine Maximum-Likelihood-Faktorenanalyse durchgeführt wurden. Bei diesen Faktorenanalysen ergab sich ein klares Ladungsmuster, lediglich bei der der Maximum-Likelihood-Faktorenanalyse zeigte sich beim Faktor *Soziale Komponente* des Wissens eine Faktorladung größer als eins, was aber bei der verwendeten Promax-Rotation auftreten kann. Beide Faktorenlösungen führten bei einer Modellierung durch ein E/CFA-Modell zu guten Fitindizes. Auch konnten einige der durch die Modifikationsindizes in das CFA-Modell eingeführten Nebenladungen in den Faktorladungen der explorativen Faktorenanalysen gefunden werden.

Man kann also festhalten, dass man die Faktorenstruktur sowohl unter der konfirmatorischen als auch der explorativen Faktorenanalyse nachweisen kann. Allerdings wurde das Gesamtmodell nicht mithilfe der Validierungsstichprobe kreuzvalidiert. Auf dieses Vorgehen wurde verzichtet, da nicht alle Faktoren im Einzelnen betrachtet einen guten Modellfit in der Validierungsstichprobe aufwiesen. Das Gesamtmodell kann also nicht über die gegebene Analysestichprobe hinaus verallgemeinert werden.

Ein Problem in dem Gesamtmodell stellen die in das Modell eingeführten Doppelladungen von Items dar, die einen Hinweis darauf geben, dass ein Item von mehreren latenten Variablen beeinflusst wird. Doppelladungen verletzen die zuvor geforderte Eindimensionalität, was

bedeutet, dass die manifesten Variablen nur durch eine latente Variable bedingt werden. In Bezug auf die Faktorenstruktur in einem CFA-Modell bedeutet Eindimensionalität, dass die Faktorladungsmatrix Λ eine Einfachstruktur aufweisen muss (Steenkamp & van Trijp, 1991). Die im Gesamtmodell hinzugefügten Nebenladungen verletzen diese Einfachstruktur. Da allerdings nur wenige der Doppelladungen signifikant sind und die Werte der numerischen standardisierten Regressionskoeffizienten eher gering ausfallen, können diese Doppelladungen für die Beurteilung der Eindimensionalität vernachlässigt werden (Marsh, 2007).

Da im weiteren Verlauf der Einfluss epistemologischer Überzeugungen auf Lernstrategien und damit einhergehend die Konstruktvalidität der EPI-Faktoren untersucht werden soll, bleibt die Frage offen, wie mit den in das Modell eingebrachten Doppelladungen umgegangen wird. Bei dieser regressionsanalytischen Fragestellung sind die epistemologischen Überzeugungen die exogenen Variablen eines Strukturgleichungsmodells. Das heißt, bis auf die ausgeschlossenen Faktoren *Geschlechtsspezifische Wissenszugänge* und *Kulturspezifische Wissenszugänge* gehen alle Faktoren und damit das evaluierte Gesamtmodell in das Strukturgleichungsmodell als exogene Variablen ein. Die endogene Variable ist jeweils eine Lernstrategie. Es hat sich gezeigt, dass die Modellpassung negativ beeinflusst wird, wenn Effekte wie Doppelladungen und Methodenfaktoren vernachlässigt werden (Kennedy, 2007). Die Vernachlässigung möglicher Doppelladungen in einem Strukturgleichungsmodell kann auch dazu führen, dass die Korrelationen der Faktoren zu hoch geschätzt und die strukturellen Beziehungen zu anderen Variablen verzerrt werden (Asparouhov & Muthén, 2009). Um diese Effekte zu vermeiden und der Gefahr vorzubeugen, dass ein Regressionsmodell, das strukturell richtig ist, aufgrund nicht modellierter Doppelladungen von Items auf mehreren Faktoren als nicht passend verworfen wird, wird für die Regression der epistemologischen Faktoren auf die Lernstrategien das Gesamtmodell mit den bereits beschriebenen Doppelladungen verwendet.

An dieser Stelle kann eine Bewertung der zweiten und dritten Forschungsfragen den EPI-Fragebogen betreffend erfolgen. Die faktorielle Validität der meisten EPI-Skalen sowie die geforderte Freiheit von korrelierenden Messfehlern und damit einhergehenden Methodenfaktoren konnte nicht nachgewiesen werden. Stattdessen waren Modifikationen der ursprünglichen Skalen notwendig. Ein problematischer Aspekt ist, dass nicht für alle Skalen die durchgeführten Modifikationen mithilfe einer Validierungsstichprobe bestätigt werden konnten. Es ist hier also zweierlei Forschungsbedarf vorhanden. Einerseits sollten die Skalen mithilfe neuer Items verbessert werden und andererseits sollte die Replizierbarkeit der faktoriellen Struk-

tur ein Ziel der weiteren Forschung sein, um eigenständig nutzbare Skalen zu erhalten. Allerdings ist auch die Konvergenz mit den Ergebnissen von Moschner und Gruber (im Druck) festzustellen, vor allem was die Teilung des Faktors *Absolutes Wissens* in zwei Faktoren betrifft. Zum jetzigen Zeitpunkt sollten die Skalen wegen der gefundenen Probleme nicht zur Diagnostik epistemologischer Überzeugungen verwendet werden.

Die Konstruktion von Skalen kann man als einen mehrstufigen Prozess beschreiben (Hildebrandt & Temme, 2006). Im ersten Schritt wird die theoretische Fundierung der zu messenden Konstrukte gelegt, im zweiten Schritt wird der Prozess mit der Generierung geeigneter Items fortgesetzt und im dritten Schritt erfolgt die Beurteilung der Anscheinsvalidität der generierten Items. Der vierte Schritt besteht aus einem Unterprozess, der sich aus wiederholten Prüfungen der entwickelten Skalen anhand statistischer Kriterien und evtl. erforderlichen Revisionen der Skala bzw. der Items zusammensetzt. Im fünften Schritt wird die auf diese Weise optimierte Skala an neuen Stichproben validiert. Verortet man die Entwicklung des EPI-Fragebogens in diesem Prozess, so befindet sich diese Entwicklung zurzeit im vierten Schritt. Bevor die Skalen nicht endgültig optimiert und abschließend validiert sind, sollte daher auf einen Einsatz in der Diagnostik epistemologischer Überzeugungen verzichtet werden.

Die vierte Forschungsfrage bezog sich auf die psychometrischen Eigenschaften der gebildeten Skalen zur Erfassung epistemologischer Überzeugungen. Dabei wurde festgestellt, dass die Reliabilität je nach Äquivalenz der durch die Skalenreliabilität ρ oder Cronbachs α ausgedrückten Messung im Allgemeinen keine zufriedenstellenden Werte erreichte. Obwohl hier schon eine Höhe der Reliabilität von .70 als ausreichend erachtet wurde (im Vergleich zu .80 wie allgemein üblich, vgl. Fisseni, 2004), konnten doch nicht alle Skalen den geforderten Wert erreichen. In Tabelle 95 sind die Reliabilitäten der Skalen zusammengefasst.

Tabelle 95: Übersicht über die Reliabilitäten der EPI-Skalen

Skala	Koeffizient	Wert
<i>Reflexive Natur des Wissens</i>	Cronbachs α	.73
<i>Absolutheit des Wissens</i>	Cronbachs α	.67
<i>Sicherheit/Objektivität des Wissens</i>	Cronbachs α	.53
<i>Soziale Komponente des Wissens</i>	Cronbachs α	.49
<i>Value of Knowledge</i>	Skalenreliabilität ρ	.64
<i>Lernen lernen</i>	Skalenreliabilität ρ	.66

Die Übersicht zeigt, dass nur beim Faktor *Reflexive Natur des Wissens* ein Wert erreicht wird, der über dem gesetzten Grenzwert von .70 liegt. Der Faktor *Lernen lernen* würde eine Relia-

bilität von .72 aufweisen, würde wie zuvor diskutiert Item 02 aus der Skala entfernt werden. Bei diesem Faktor liegt die Vermutung zugrunde, dass es sich ähnlich den Gegebenheiten beim Schommer-Fragebogen um einen Faktor mit einer hierarchischen Struktur handeln könnte. Allerdings lässt sich bei keinem anderen Faktor auf der Grundlage inhaltlicher Aspekte der Items die Vermutung einer hierarchischen bzw. mehrdimensionalen Struktur anstellen. Es bleibt also die Frage offen, warum die meisten der Skalen solch eine geringe Reliabilität besitzen.

Moschner und Gruber (2005b) nehmen an, dass epistemologische Überzeugungen als States und nicht als Traits zu konzeptualisieren sind. Ein State ist ein zeitlich instabiles, zustands- und situationsabhängiges Merkmal, wohingegen ein Trait ein zeitlich stabiles, zustands- und situationsunabhängiges Merkmal darstellt (Kelava & Schermelleh-Engel, 2008). Epistemologische Überzeugungen sind erfahrungs- und wissensabhängig und können sich verändern, sie sind somit in den Kontext desjenigen Wissens eingebettet, auf welches sie sich beziehen. Entsprechend sind Personen- und Situationsvariablen sowie die Interaktion von Person und Situation verschachtelt. Nach der Latent-State-Trait-Theorie (Steyer, Ferring & Schmitt, 1992, Steyer, Schmitt & Eid, 1999) gibt die Personenvariable Auskunft über den Trait-Anteil epistemologischer Überzeugungen, während die Situation und die Interaktion von Person und Situation den State-Anteil epistemologischer Überzeugungen beschreiben. Im Rahmen der LST-Theorie lässt sich die Varianz einer Variablen in Anteile zerlegen, welche die Stabilität und Variabilität angeben. Die Stabilität drückt dabei den Trait-Anteil aus, die Variabilität drückt den State-Anteil der Variablen aus. Empirisch wäre nun zu überprüfen, wie groß der Anteil der Stabilität und wie groß der Anteil der Variabilität an der Gesamtvarianz einer Variablen ist. Würde epistemologische Überzeugungen einen State darstellen, so würde dies dazu führend, dass die Reliabilität sinkt, da sich diese auf den Trait-Anteil bezieht. Diese Möglichkeit wird in Kapiteln 6.3.3 und 6.5 ausführlich diskutiert.

Es ist also die Frage aufgeworfen worden, ob die in Kapitel 1.7 skizzierte Sichtweise epistemologischer Überzeugungen als Traits gerechtfertigt ist. Die Annahme von Traits wurde schon von Hammer und Elby (2002, vgl. Kapitel 1.4) kritisiert, wobei sich diese Kritik hauptsächlich auf die Abgrenzung des von Hammer und Elby propagierten Modells epistemologischer Ressourcen gegenüber den herkömmlichen persönlichkeitspsychologischen Ansätzen bezieht. Elby und Hammer (2001) haben aber ebenso die Verwendung der üblichen Fragebogenstatements zur Erfassung epistemologischer Überzeugungen und die übliche normative

Sichtweise fortgeschrittener epistemologischer Überzeugungen kritisiert, die übliche Erfassung als Traits klammert erforderliche Anpassungen an die Lernsituation aus. Die Konzeption epistemologischer Überzeugungen mittels der LST-Theorie wird bei der abschließenden Diskussion ausführlicher besprochen.

In Bezug auf die psychometrischen Eigenschaften der entwickelten Skalen kann festgehalten werden, dass eine Erfassung epistemologischer Überzeugungen möglich, diese Erfassung aber nicht reliabel ist. Es besteht weiterer Forschungsbedarf, um zu klären, ob es sich bei epistemologischen Überzeugungen um States oder um Traits handelt. An dieser Stelle kann die zuvor gegebene Empfehlung, die EPI-Skalen vorerst nicht zur Diagnostik epistemologischer Überzeugungen zu verwenden, um den Aspekt der mangelnden Reliabilität dieser Skalen erweitert werden.

Bis hier wurden die Aspekte der Konstruktvalidität auf Itemebene bzw. der faktoriellen Validität behandelt. Eine wesentliche Forschungsfrage betrifft die Relevanz epistemologischer Überzeugungen für Lernprozesse, d. h. ob epistemologische Überzeugungen den eigentlichen Prozess des Wissenserwerbs beeinflussen. Die Frage zielt auf die konkurrente bzw. prädikative Validität epistemologischer Überzeugungen ab, was im nächsten Kapitel behandelt wird.

5.3 Konstruktvalidierung des EPI

5.3.1 Faktorenstruktur des LIST

In diesem Kapitel soll die fünfte Forschungsfrage zum EPI-Fragebogen beantwortet werden, das ist die Frage nach dem Zusammenhang der EPI-Faktoren mit Lernstrategien. Hierzu ist es allerdings erforderlich, Lernstrategien mithilfe eines geeigneten Instruments zu erfassen. Dieses ist das Inventar zur Erfassung von Lernstrategien im Studium (LIST, s. Anhang D für eine Übersicht über die Items des LIST). Da wie in Kapitel 1.6.2 erwähnt der LIST selbst nicht ausreichend auf seine faktorielle Validität hin untersucht wurde, wird vor den eigentlichen Regressionsanalysen eine Überprüfung der LIST Skalen mithilfe konfirmatorischer Faktorenanalysen durchgeführt. Es schließen sich die Regressionen epistemologischer Überzeugungen auf Lernstrategien und die Diskussion der Ergebnisse an.

5.3.1.1 Vorbereitung der Daten

Ausgehend von dem oben genannten Datensatz von 419 Lehramtsstudierenden der Universität des Saarlandes wurden diejenigen Fälle ausgeschlossen, die fünf oder mehr fehlende Werte bei den Items des EPI und sieben oder mehr fehlende Werte bei den Items des LIST hatten. Die Annahme ist auch hier, dass mehr als 10 % fehlende Werte auf ein willkürliches Antwortverhalten hindeuten. Dies führte zum Ausschluss von 174 Fällen, sodass die Daten von 245 Versuchspersonen übrig blieben. Bei den verbleibenden 245 Fällen wurden wie zuvor die fehlenden Werte mittels des EM-Algorithmus geschätzt (Allison, 2012, 2009, Lüdtke et al., 2007). Die Altersverteilung der Stichprobe ist in Tabelle 96, die Aufteilung der Probanden auf die Fächergruppen ist in Tabelle 97 zu finden.

Tabelle 96: Deskriptive Statistik der Altersverteilung in Jahren

Stichprobe	<i>n</i>	Range	Min	Max	<i>M</i>	<i>SD</i>
männlich	98	22	19	41	22.96	4.47
weiblich	147	29	18	47	21.04	3.80
gesamt	245	29	18	47	21.81	4.18

Tabelle 97: Aufteilung der Probanden auf die Fächergruppen

Fach	Geslecht		männlich		weiblich	
	Häufigkeit	%	Häufigkeit	%	Häufigkeit	%
Wirtschaftspädagogik/BWL	14	5.7	9	9.2	5	3.4
Germanistik	50	20.4	14	14.3	36	24.5
Sprachen	68	27.8	17	17.3	51	34.7
Sport	15	6.1	7	7.1	8	5.4
Naturwissenschaften	59	24.1	35	35.7	24	16.3
Gesellschafts-/Kulturwissenschaften	29	11.8	12	12.2	17	11.6
Musik/Kunst	5	2.0	1	1.0	4	2.7
Keine Angabe	5	2.0	3	3.1	2	1.4
Gesamt	245	100	98	100	147	100

Die 77 Items des LIST weisen eine Abweichung von der multivariaten Normalverteilung auf, der Mardia-Test liefert ein signifikantes Ergebnis ($M = 455.88$, $CR = 32.35$). Sonst zeigt keines der Items einen Wert über den Grenzen von West et al. (1995), also eine Schiefe > 2 oder eine Wölbung > 7 . Durch die teilweise vorhandenen Schiefen einiger Itemverteilungen und wegen des signifikanten Mardia-Tests kann nicht ausgeschlossen werden, dass die Parameterschätzungen verzerrt sind. Daher wird standardmäßig die Bollen-Stine-Bootstrap-Methode angewandt und die Möglichkeit von Verzerrung bei der Schätzung von Parametern wird bei der Interpretation der Faktorladungen berücksichtigt.⁵¹

Tabelle 98: Verteilungsform der LIST Items

Variable	Schiefe	CR	Wölbung	CR
Item 01	.08	.49	-.97	-3.11
Item 02	-.50	-3.17	-.18	-.57
Item 03	-.57	-3.62	-.02	-.07
Item 04	-1.18	-7.54	1.02	3.25
Item 05	-.71	-4.53	.09	.30
Item 06	-1.27	-8.11	1.85	5.90
Item 07	-.07	-.45	-.55	-1.77
Item 08	-.51	-3.23	-.23	-.73
Item 09	-1.00	-6.39	1.15	3.68
Item 10	.09	.60	-.05	-.15
Item 11	-.22	-1.43	-.92	-2.93
Item 12	-.91	-5.84	.63	2.03

⁵¹ Da die Probanden bei dieser Untersuchung weitestgehend mit den Proband aus der Untersuchung zu den EPI-Skalen identisch sind und da bei der Analyse der Itemverteilungen in Kapitel 5.1.1 fast keine merkliche Abweichung von der Normalverteilung im Sinne von West et al (1995) berichteten Grenzen festgestellt wurden, wird an der Stelle auf eine gesonderte Untersuchung der Verteilungsform der EPI-Items verzichtet. Unterschiede zwischen den benutzten Datensätzen der ergeben sich lediglich durch den Ausschluss von Probanden mit fehlenden Werten.

Konstruktvalidierung des EPI

Variable	Schiefe	CR	Wölbung	CR
Item 13	-.91	-5.83	1.03	3.29
Item 14	-.71	-4.51	-.33	-1.04
Item 15	-.04	-.26	-.49	-1.58
Item 16	-.84	-5.34	-.16	-.50
Item 17	.07	.42	-.59	-1.90
Item 18	.32	2.05	-.36	-1.16
Item 19	-.51	-3.28	-.01	-.02
Item 20	-.59	-3.74	-.50	-1.61
Item 21	-.82	-5.25	.62	1.98
Item 22	.42	2.67	-.08	-.24
Item 23	-.40	-2.57	-.98	-3.14
Item 24	-.81	-5.17	-.28	-.88
Item 25	-.47	-3.01	-.57	-1.82
Item 26	.04	.23	-.69	-2.21
Item 27	-.71	-4.55	.47	1.52
Item 28	-.12	-.76	-.40	-1.28
Item 29	.20	1.29	-.30	-.95
Item 30	-.80	-5.14	.17	.54
Item 31	-.61	-3.88	-.27	-.86
Item 32	-.43	-2.75	-.46	-1.47
Item 33	.01	.08	-.49	-1.55
Item 34	.04	.26	-1.01	-3.22
Item 35	-.50	-3.19	-.75	-2.40
Item 36	-.65	-4.17	-.20	-.64
Item 37	.15	.97	-.63	-2.02
Item 38	-.78	-4.99	.14	.46
Item 39	-.57	-3.64	-.42	-1.33
Item 40	.00	-.02	-.27	-.86
Item 41	-.09	-.57	-.85	-2.71
Item 42	-.93	-5.91	.73	2.33
Item 43	-1.15	-7.38	.88	2.81
Item 44	.33	2.13	-.21	-.67
Item 45	-.04	-.27	-.92	-2.94
Item 46	-1.10	-7.04	.96	3.06
Item 47	-.36	-2.32	-.68	-2.16
Item 48	-.64	-4.06	-.07	-.23
Item 49	-.40	-2.58	-.20	-.62
Item 50	-.02	-.15	-.56	-1.79
Item 51	.79	5.06	-.26	-.83
Item 52	-.44	-2.84	-.40	-1.29
Item 53	-.16	-1.01	-.70	-2.24
Item 54	.07	.44	-.37	-1.19
Item 55	-.69	-4.43	-.32	-1.02
Item 56	-1.17	-7.45	.90	2.87
Item 57	-.59	-3.79	-.26	-.83
Item 58	-.09	-.55	-.23	-.72
Item 59	.05	.30	-.12	-.40
Item 60	-.59	-3.79	-.41	-1.31
Item 61	-.33	-2.12	-.38	-1.22
Item 62	-.91	-5.79	.37	1.17

Variable	Schiefe	CR	Wölbung	CR
Item 63	.36	2.31	-.42	-1.34
Item 64	-.84	-5.36	.30	.95
Item 65	-.39	-2.49	-.60	-1.92
Item 66	-.63	-4.02	-.28	-.90
Item 67	-.05	-.32	-.19	-.61
Item 68	.16	1.04	-.43	-1.37
Item 69	.06	.39	-.84	-2.67
Item 70	-.24	-1.55	-.44	-1.41
Item 71	.18	1.16	-.32	-1.02
Item 72	-.37	-2.38	-.70	-2.24
Item 73	-.24	-1.53	-.50	-1.61
Item 74	-.03	-.18	-.54	-1.73
Item 75	-.08	-.54	-.58	-1.84
Item 76	-.90	-5.73	.83	2.66
Item 77	-1.17	-7.49	1.77	5.66
Mardia-Test			455.88	32.35

5.3.1.2 Konfirmatorische Faktorenanalysen des LIST

Bevor die Konstruktvalidität der EPI-Faktoren mittels des Zusammenhangs mit den Lernstrategien des LIST evaluiert wird, soll eine konfirmatorische Faktorenanalyse der einzelnen List-Skalen durchgeführt werden. Die CFA dient einerseits dazu sicherzustellen, dass die LIST Skalen dem direkt-reflexiven Messmodell (Edwards & Bagozzi, 2000) genügen und eindimensional (d. h. ohne korrelierende Messfehler) sind. Würden die Skalen unüberprüft in ein Strukturgleichungsmodell aufgenommen, bestünde die Gefahr, dass wegen eines unpassenden Messmodells eine schlechte Modellanpassung aufgrund lokaler Modellfehlspezifikationen entstehen würde. Dies hätte zur Folge, dass die Beziehung der epistemologischen Überzeugungen und der Lernstrategie nicht beurteilt werden könnte, da das Modell als nicht passend zurückgewiesen worden wäre. Überdies muss vor einer Anwendung in einem Strukturgleichungsmodell die faktorielle Struktur der Items zur Erfassung der metakognitiven Strategien geklärt werden. Ursprünglich wurden von Wild und Schiefele (1994) in diesem Bereich die drei Strategien *Planung*, *Regulation* und *Selbstüberwachung* angenommen, die in der anschließenden Analyse faktoriell nicht getrennt werden konnten und zu der Strategie *Metakognitive Strategie* zusammengefasst wurden. Allerdings gibt es Hinweise, dass diese drei Strategien doch in drei Faktoren aufgeteilt werden können (Boerner et al., 2005).

Die konfirmatorischen Faktorenanalysen der LIST-Skalen haben ein anderes Ziel als die der EPI-Skalen. Während bei der Überprüfung der EPI-Skalen eine ggf. erforderliche Revision des Messmodells im Vordergrund stand, soll bei den LIST-Skalen – bis auf Ausnahme der

metakognitiven Strategien – möglichst keine Modifikation vorgenommen werden. Lediglich einzelne Items sollten bei einer Modellfehlspezifikation ausgeschlossen werden, wobei der Ausschluss nur dann erfolgen soll, wenn die Modellpassung massiv beeinflusst wird. Auf eine Evaluation der endgültigen Messmodelle der LIST-Skalen wird an dieser Stelle verzichtet, da dies außerhalb der Ziele der vorliegenden Arbeit liegt.

5.3.1.2.1 CFA der Strategie Elaborieren

Das initiale Messmodell zeigte einen guten bis akzeptablen Modellfit ($\chi^2 = 36.66$ ($df = 20$), $p = .01$, Bollen-Stine- $p = .09$, $\chi^2/df = 1.83$, CFI = .97, TLI = .96, SRMR = .0384, RMSEA = .058, CI = [.028;.088], $p_{close} = .29$). Die Inspektion der Residuen und Modifikationsindizes ergab, dass zwischen Item 28 und Item 49 möglicherweise ein korrelierender Fehler existiert. Beide Items beziehen sich inhaltlich auf den gleichen Sachverhalt und haben die Formulierung „Ich versuche ...“ gemeinsam. Es wurde Item 28 ausgeschlossen, da dessen Formulierung im Vergleich mit Item 48 umständlich erscheint.

Das resultierende Messmodell wies einen guten Modellfit auf ($\chi^2 = 18.28$ ($df = 14$), $p = .19$, Bollen-Stine- $p = .41$, $\chi^2/df = 1.30$, CFI = .99, TLI = .98, SRMR = .0307, RMSEA = .035, CI = [0;.076], $p_{close} = .68$). Das für die weitere Analyse benutzte Messmodell der Strategie *Elaborieren* besteht also aus den Items: 02, 17, 39, 49, 58, 67, 73. Cronbachs α der Skala *Elaborieren* beträgt .83, die Skalenreliabilität beträgt ebenfalls .83.

5.3.1.2.2 CFA der Strategie Kritisches Prüfen

Das initiale Messmodell besaß eine gute bis akzeptable Modellpassung ($\chi^2 = 37.70$ ($df = 20$), $p = .01$, Bollen-Stine- $p = .08$, $\chi^2/df = 1.89$, CFI = .97, TLI = .96, SRMR = .0395, RMSEA = .060, CI = [.029;.089], $p_{close} = .26$). Die Residuen und die Modifikationsindizes deuten auf einen korrelierenden Messfehler zwischen den Items 03 und 59 hin. Inhaltlich sind die Items gleichwertig. Da die Formulierung von Item 03 deutlich schwieriger ist als die von Item 59 und auch die Faktorladung von Item 03 geringer ist als die von Item 59, wurde das Item 03 aus der weiteren Analyse entfernt.

Das resultierende Modell zeigte beim RMSEA und dem p-Wert eine akzeptable Modellpassung, die übrigen Indizes deuteten auf eine gute Modellpassung ($\chi^2 = 23.91$ ($df = 14$), $p = .05$, Bollen-Stine- $p = .16$, $\chi^2/df = 1.70$, CFI = .98, TLI = .98, SRMR = .0312, RMSEA = .054, CI = [.006;.090], $p_{close} = .39$). Durch die Residuen und Modifikationsindizes ergaben sich kei-

ne Hinweise auf eine weitere Fehlspezifikation. Das endgültige Messmodell für die Strategie *Kritisches Prüfen* besteht aus den Items 18, 29, 40, 50, 59, 68 und 74. Cronbachs α der Skala *Kritisches Prüfen* beträgt .86, auch die Skalenreliabilität beträgt .86.

5.3.1.2.3 CFA der Strategie Organisieren

Das ursprüngliche Modell wies, mit Ausnahme des CFI und TLI, nur eine akzeptable Modellpassung auf ($\chi^2 = 44.10$ ($df = 20$), $p = .001$, Bollen-Stine- $p = .02$, $\chi^2/df = 2.20$, CFI = .96, TLI = .95, SRMR = .0479, RMSEA = .070, CI = [.042;.098], $p_{close} = .10$). Die Inspektion der Residuen deutet auf Probleme mit dem Item 56 hin. Eine Überprüfung des Inhaltes ergab, dass sich das Item nicht auf eine Reorganisation der Lerninhalte bezieht, die durch die Items erfasst werden soll. Deshalb wurde das Item ausgeschlossen.

Der Fit des resultierenden Modells verbesserte sich deutlich ($\chi^2 = 23.44$ ($df = 14$), $p = 0.05$, Bollen-Stine- $p = .12$, $\chi^2/df = 1.67$, CFI = .99, TLI = .98, SRMR = .0347, RMSEA = .053, CI = [0;.089], $p_{close} = .41$). Obwohl das Modell teilweise nur im Bereich akzeptabler Modellpassung liegt, wird keine weitere Modifikation vorgenommen, um so weit wie möglich an der ursprünglichen Skala zu bleiben. Das im weiteren Verlauf benutzte Messmodell für die Strategie *Organisation* besteht aus den Items 01, 14, 25, 36, 47, 65 und 72. Cronbachs α der Skala *Organisieren* beträgt .84, die Skalenreliabilität beträgt .85.

5.3.1.2.4 CFA der Strategie Wiederholen

Das ursprüngliche Modell lag überwiegend im Bereich akzeptabler Modellpassung ($\chi^2 = 30.86$ ($df = 14$), $p = .006$, Bollen-Stine- $p = .05$, $\chi^2/df = 2.20$, CFI = .95, TLI = .93, SRMR = .0474, RMSEA = .070, CI = [.036;.104], $p_{close} = .15$). Die Inspektion der Residuen und der Modifikationsindizes deuten darauf hin, dass zwischen Item 51 und 69 wahrscheinlich ein korrelierender Fehler existiert. Inhaltlich sind die Items 51 und 69 ähnlich, beide beziehen sich auf Auswendiglernen. Da das Item 51 die niedrigere Faktorladung auswies, wurde dieses aus der weiteren Analyse ausgeschlossen.

Das resultierende Modell führte zu einer guten Modellanpassung ($\chi^2 = 12.80$ ($df = 9$), $p = .17$, Bollen-Stine- $p = .31$, $\chi^2/df = 1.42$, CFI = .99, TLI = .98, SRMR = .0364, RMSEA = .042, CI = [0;.089], $p_{close} = .56$). Das im Weiteren benutzte Messmodell für diese Strategie besteht aus den Items 04, 19, 30, 41, 60 und 69. Cronbachs α der Skala *Wiederholen* beträgt .76, die Skalenreliabilität beträgt .77.

5.3.1.2.5 Analyse der metakognitiven Strategien

Die ursprüngliche vorgesehene Zuordnung der Items zu jeder der drei metakognitiven Strategien von Wild und Schiefele (1994) war:

- *Planen*: 05, 20, 31,42,
- *Überwachen*: 52, 61,70, 75,
- *Regulieren*: 06, 13, 77.

Diese Zuordnung ließ sich bei Wild und Schiefele (1994) mittels der durchgeführten Hauptkomponentenanalysen nicht bestätigen, sodass in der Folge alle Items zu einer Skala *Metakognitive Strategien* zusammengefasst wurden. Diese Zusammenfassung ergibt einen inakzeptablen Modellfit ($\chi^2 = 115.63$ ($df = 44$), $p < .01$, Bollen-Stine- $p = .001$, $\chi^2/df = 2.63$, CFI = .85, TLI = .80, SRMR = .0664, RMSEA = .082, CI = [.064;.100], $p_{close} = .003$). Deswegen wurde mithilfe einer EFA überprüft, ob sich eine theoretisch angenommene 3-faktorielle Struktur identifizieren lässt. Dazu wurden mit allen Items, welche die metakognitiven Strategien erfassen sollen, eine Hauptachsenanalyse mit anschließender Promax-Rotation ($\kappa = 4$) durchgeführt. Das KMO-Maß betrug .82. Obwohl der Scree-Plot nicht auf eine 3-faktorielle Struktur hindeutete, zeigte sich in der Mustermatrix eine klare 3-faktorielle Struktur (s. Tabelle 99). Wie nicht anders zu erwarten, korrelierten die resultierenden Faktoren sehr stark (s. Tabelle 100).

Tabelle 99: EFA der metakognitiven Strategien

Item	Faktor		
	Planen	Überwachen	Regulieren
Item 31	.66	.21	-.17
Item 42	.61	-.07	.15
Item 20	.56	.04	-.06
Item 05	.35	-.15	.10
Item 52	-.07	.63	.10
Item 75	-.05	.58	.00
Item 70	.18	.51	.03
Item 77	.22	.01	.62
Item 61	-.25	.21	.53
Item 06	.12	-.07	.45
Item 13	.26	.00	.36

Tabelle 100: Korrelation der drei Faktoren

Faktor	Planen	Überwachen	Regulieren
Planen	-		
Überwachen	.51	-	
Regulieren	.54	.48	-

Die Mustermatrix deutet auf eine klare 3-faktorielle Struktur hin, es existieren keine substantiellen Nebenladungen. Auf dem ersten Faktor laden die Items 31, 42, 20 und 05, dieser Faktor lässt sich als die Strategie *Planen* identifizieren. Auf dem zweiten Faktor laden die Items 52, 75 und 70. Dies sind bis auf Item 61 die Items der Strategie *Überwachen*. Auf dem dritten Faktor laden die Items 77, 61, 06 und 13, dies sind alle Items der Strategie *Regulieren* und zusätzlich das Item 61, das ursprünglich der Strategie *Überwachen* zugeordnet war. Eine Inspektion von Item 61 ergab, dass es im Sinne eine Regulation verstanden werden kann. Im Folgenden wird das Item deshalb dem Faktor *Regulation* zugeordnet. Der zweite Faktor lässt sich somit als die Strategie *Überwachen* und der dritte Faktor als die Strategie *Regulieren* identifizieren.

Auf der Grundlage dieser EFA wurde ein CFA-Modell mit drei latenten Variablen aufgestellt, dabei wurden die Items mit Hauptladungen auf den EFA-Faktoren als manifeste Variable in der CFA aufgenommen. Die Zuordnung der Items zu den Faktoren ist für das CFA-Modell wie folgt: Die Strategie *Planen* umfasst die Items 05, 20, 31 und 42, die Strategie *Überwachen* umfasst die Items 52, 70 und 75, die Strategie *Regulieren* besteht aus den Items 06, 13, 77 und 61.

Doppelladungen und korrelierende Fehler wurden nicht in das Modell aufgenommen, es wurden alle drei möglichen Korrelationen der drei latenten Variablen frei geschätzt. Das Modell führte zu einer guten bis akzeptablen Modellpassung ($\chi^2 = 61.80$ ($df = 41$), $p = .02$, Bollen-Stine- $p = .17$, $\chi^2/df = 1.50$, CFI = .96, TLI = .93, SRMR = .0486, RMSEA = .046, CI = [.019;.068], $p_{close} = .60$).

Cronbachs α der Skala *Planen* beträgt .67, der Skala *Überwachen* .63 und der Skala *Regulieren* .62. Die Skalenreliabilitäten betragen entsprechend .64, .63 und .63, wobei bei der Berechnung der Skalenreliabilität der Skala *Überwachen* aus Gründen der Modellidentifikation von der τ -Äquivalenz der Items ausgegangen werden musste. Für die in Kapitel 5.3.2 folgenden Analysen werden die Strategien *Planen*, *Überwachen* und *Regulieren* jeweils als einzelnen Faktoren betrachtet, d. h. diese gehen separat in das Strukturgleichungsmodell ein.

5.3.1.2.6 CFA der Strategie Anstrengung

Das initiale Modell zeigte gute bis akzeptable Fitindizes ($\chi^2 = 37.87$ ($df = 20$), $p = .01$, Bollen-Stine- $p = .13$, $\chi^2/df = 1.89$, CFI = .96, TLI = .94, SRMR = .0447, RMSEA = .061

(CI = [.030;.090], $p_{close} = .26$). Eine Inspektion der Residuen ergab, dass sich die Korrelation von Item 43 und 53 durch das Modell nicht replizieren lässt. Inhaltlich passt Item 43 nicht eindeutig zu dem Rest der Items, da es auch im Sinne von Zeitmanagement interpretiert werden kann.

Schließt man Item 43 aus, weisen bis auf den p-Wert alle Fitindizes eine gute Modellpassung auf ($\chi^2 = 21.84$ ($df = 14$), $p = .08$, Bollen-Stine- $p = .37$, $\chi^2/df = 1.56$, CFI = .98, TLI = .95, SRMR = .0330, RMSEA = .048, CI = [.000;.085], $p_{close} = .49$). Das im Folgende benutzte Messmodell für die Strategie *Anstrengung* besteht aus den Items 09, 21, 32, 53, 62, 71, 76. Cronbachs α der Skala *Anstrengung* beträgt .81, die Skalenreliabilität beträgt ebenfalls .81.

5.3.1.2.7 CFA der Strategie Aufmerksamkeit

Das ursprüngliche Modell zeigte sowohl gute als auch nur akzeptable Fitindizes ($\chi^2 = 22.25$ ($df = 9$), $p = .008$, Bollen-Stine- $p = .09$, $\chi^2/df = 2.47$, CFI = .99, TLI = .98, SRMR = .0256, RMSEA = .078, CI = [.037;.119], $p_{close} = .12$). Die Residuen gaben keinen Hinweis auf evtl. vorliegende Fehlspezifikationen. Es gab zwar Modifikationsindizes mit einem Betrag von größer als 4, deren geschätzte Parameteränderung war aber nur minimal, aus statistischen Gründen war somit keine Modelländerung naheliegend. Cronbachs α der Skala *Aufmerksamkeit* beträgt .91, die Skalenreliabilität beträgt .91.

5.3.1.2.8 CFA der Strategie Zeitmanagement

Die Skala zur Erfassung der Strategie Zeitmanagement wies eine gute Modellpassung auf ($\chi^2 = .66$ ($df = 2$), $p = .73$, Bollen-Stine- $p = .74$, $\chi^2/df = .33$, CFI = 1.00, TLI = 1.00, SRMR = .0063, RMSEA = .000, CI = [.000;.091], $p_{close} = .83$). Lediglich das Konfidenzintervall des RMSEA war sehr breit. Da die Residuen und die Modifikationsindizes keinen Hinweis auf eine etwa vorliegende Fehlspezifikation gaben, wurde die Skala unverändert übernommen. Cronbachs α der Skala Zeitmanagement beträgt .87, die Skalenreliabilität beträgt ebenfalls .87.

5.3.1.2.9 CFA der Strategie Lernumgebung

Das initiale Modell ergab keine akzeptable Modellpassung ($\chi^2 = 93.76$ ($df = 9$), $p < .01$, Bollen-Stine- $p = .001$, $\chi^2/df = 10.42$, CFI = .80, TLI = .66, SRMR = .086, RMSEA = .196, CI = [.162;.233], $p_{close} < .001$). Die Residuen deuteten auf eine massive Überschätzung der Korrelation von Item 64 und Item 55 hin. Der Modifikationsindex lag bei 76.55 und die ge-

geschätzte Korrelation der Fehlervariablen lag bei .43. Inhaltlich sind Item 64 und 55 fast identisch, beide beziehen sich auf die Ordnung am Arbeitsplatz. Da Item 64 die im Vergleich zu Item 55 niedrigere Faktorladung aufwies, wurde es ausgeschlossen.

Das revidierte Modell führte zu einer guten Modellpassung ($\chi^2 = 2.27$ ($df = 5$), $p = .81$, Bollen-Stine- $p = .84$, $\chi^2/df = .46$, CFI = 1.00, TLI = 1.02, SRMR = .0160, RMSEA = .000, CI = [0;.055], $p_{close} = .94$). Das finale Modell beinhaltet die Items 12, 24, 35, 46 und 55. Cronbachs α der Skala *Lernumgebung* beträgt .73, die Skalenreliabilität beträgt .74.

5.3.1.2.10 CFA der Strategie Lernen mit Studienkollegen

Das initiale Modell wies keine akzeptable Modellpassung auf ($\chi^2 = 119.93$ ($df = 14$), $p < .001$, Bollen-Stine- $p = .001$, $\chi^2/df = 8.57$, CFI = .83, TLI = .75, SRMR = .0892, RMSEA = .176, CI = [.148;.206], $p_{close} < .001$). Die Inspektion der Residuen ergab mehrere mögliche Fehlspezifikationen. Die Modifikationsindizes deuten auf massive Korrelationen der Fehlervariablen hin, von 21 möglichen paarweisen Fehlerkorrelationen weisen 18 einen Modifikationsindex größer als vier auf (vgl. Tabelle 101).

Tabelle 101: Modifikationsindizes für die Strategie Lernen mit Studienkollegen

Vorgeschlagener Pfad		Modifikationsindex	erwartete Parameteränderung	
Item 07	→	Item 15	26.43	.27
Item 26	→	Item 15	6.45	.13
Item 37	→	Item 15	15.52	.19
Item 15	→	Item 07	23.38	.27
Item 37	→	Item 07	5.13	.12
Item 66	→	Item 07	6.16	-.14
Item 15	→	Item 26	6.01	.14
Item 37	→	Item 26	4.84	.12
Item 15	→	Item 37	13.57	.23
Item 07	→	Item 37	5.07	.14
Item 26	→	Item 37	4.54	.12
Item 07	→	Item 48	5.19	-.10
Item 26	→	Item 48	5.44	-.10
Item 15	→	Item 57	5.27	-.10
Item 37	→	Item 57	7.90	-.11
Item 48	→	Item 57	4.66	.10
Item 15	→	Item 66	5.39	-.12
Item 07	→	Item 66	9.75	-.15

Beim Vergleich der Items zeigt sich, dass bis auf Item 48 in jedem die Formulierung Studienkollegen enthalten ist. Die möglicherweise vorhandenen Fehlerkorrelationen können also

durch eine ähnliche Formulierung der Items verursacht worden sein und wären damit als Methodenfaktor zu deuten. Da eine Revision des Modells einen massiven Eingriff in die Skala bedeutet hätte, wurde auf eine Respezifikation verzichtet und die Skala *Lernen mit Studienkollegen* in die nachfolgenden Analysen nicht aufgenommen.

5.3.1.2.11 CFA der Strategie Literatur

Das ursprüngliche Modell zeigte eine gute Modellpassung ($\chi^2 = .32$ ($df = 2$), $p = .85$, Bollen-Stine- $p = .90$, $\chi^2/df = .16$, CFI = 1.00, TLI = 1.02, SRMR = .0064, RMSEA = .000, CI = [0;.069], $p_{close} = .92$). Die Skala wurde unverändert in den weiteren Strukturgleichungsmodellen verwandt. Cronbachs α der Skala *Literatur* beträgt .76, die Skalenreliabilität beträgt .78.

5.3.1.2.12 Kurzes Fazit der konfirmatorischen Faktorenanalyse des LIST

Im Bereich metakognitiver Strategien ließen sich die drei Strategien *Regulieren*, *Planen* und *Überwachen* nachweisen, was nach der ursprünglichen Theorie von Wild und Schiefele (1994) sowie den Ergebnissen von Boerner et al. (2005) zu erwarten war. Bei den Skalen zur Erfassung der Strategie *Lernen mit Studienkollegen* ergaben sich Probleme im Hinblick auf die Modellpassung, sodass diese Skala ausgeschlossen wird.

Was nicht geprüft wurde, ist die psychometrische Äquivalenz der Skalen, da diese für die weitere Analyse nicht von Belang ist. Daher sollten nur die Skalenreliabilitäten als Reliabilitätsschätzung betrachtet werden, da die Erfüllung der Voraussetzungen für eine sinnvolle Interpretation von Cronbachs α somit nicht bekannt ist.

5.3.2 Zusammenhänge epistemologischer Überzeugungen mit Lernstrategien

Das Hauptanliegen dieser Untersuchung ist die Validierung der EPI-Skalen zur Erfassung epistemologischer Überzeugungen. Die oben durchgeführte Untersuchung der faktoriellen Validität der EPI war der erste Pfeiler der Validierung. Der zweite Pfeiler ist die Bestimmung der Konstruktvalidität der EPI-Skalen. Die Konstruktvalidierung erfolgt nach dem in Kapitel 3.6.3 beschriebenen Vorgehen zur Bestimmung der konkurrenten Validität. Berichtet werden jeweils die Modellpassung des Strukturgleichungsmodells sowie die Pfadkoeffizienten der epistemologischen Überzeugungen auf die Lernstrategien. Die Faktorladungen der manifesten auf die latenten Variablen des Mess- bzw. Strukturmodells werden nicht berichtet.

5.3.2.1 Elaborieren

Die Modellanpassung liegt im guten bis akzeptablen Bereich ($\chi^2 = 534.17$ ($df = 456$), $p = .01$, Bollen-Stine- $p = .52$, $\chi^2/df = 1.17$, CFI = .95, TLI = .94, SRMR = .0531, RMSEA = .027, CI = [.015;.035], $p_{close} = 1.00$). Der einzig signifikante Pfadkoeffizient ist zwischen dem Faktor *Reflexive Natur* des Wissens und der Lernstrategie *Elaborieren*. Die Varianzaufklärung beträgt 16.1%.

Tabelle 102: Regression EPI auf Elaborieren

Skala	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
RN	.53	.44	.16	3.22	***
AW	.16	.15	.12	1.33	.18
S/OW	-.16	-.15	.12	-1.31	.19
SK	.04	.04	.13	.29	.77
VK	-.04	-.03	.18	-.23	.82
LL	.01	.01	.21	.07	.95

Anmerkung: ***: $p < .001$

5.3.2.2 Kritisches Prüfen

Die Modellanpassung liegt im guten bis akzeptablen Bereich, lediglich der TLI liegt unter der Grenze akzeptabler Modellpassung ($\chi^2 = 620.821$ ($df = 456$), $p < .001$, Bollen-Stine- $p = .154$, $\chi^2/df = 1.36$, CFI = .90, TLI = .89, SRMR = .0573, RMSEA = .038, CI = [.031;.046], $p_{close} = .99$). Auch hier ist der einzig signifikante Pfadkoeffizient zwischen dem Faktor *Reflexive Natur des Wissens* und der Strategie *Kritisches Prüfen*, die Varianzaufklärung des Modells beträgt 15.2 %.

Tabelle 103: Regression EPI auf Kritisches Prüfen

Skala	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
RN	.56	.42	.18	3.15	***
AW	-.02	-.02	.14	-.14	.89
S/OW	.01	.01	.14	.10	.92
SK	.23	.22	.16	1.44	.15
VK	-.36	-.19	.24	-1.51	.13
LL	.08	.04	.23	.36	.72

Anmerkung: ***: $p < .001$

5.3.2.3 Organisieren

Die Modellanpassung liegt im akzeptablen Bereich ($\chi^2 = 550.66$ ($df = 456$), $p = .002$, Bollen-Stine- $p = .44$, $\chi^2/df = 1.20$, CFI = .94, TLI = .94, SRMR = .0526, RMSEA = .029, CI = [.019;.038], $p_{close} = 1.00$). Die Strategie *Organisieren* wird nur durch den Faktor *Lernen lernen* wesentlich vorhergesagt, alle anderen Pfadkoeffizienten sind nicht signifikant. Alle anderen epistemologischen Überzeugungen zeigen keinen signifikanten Einfluss auf diese Lernstrategie. Die Varianzaufklärung beträgt 15.6 %.

Tabelle 104: Regression EPI auf Organisation

Skala	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
RN	.10	.08	.15	.67	.50
AW	.14	.13	.11	1.25	.21
S/OW	-.13	-.12	.12	-1.11	.27
SK	.06	.07	.13	.48	.63
VK	-.12	-.07	.18	-.68	.50
LL	.71	.35	.28	2.56	.01

5.3.2.4 Wiederholen

Die Modellpassung liegt im guten bis akzeptablen Bereich ($\chi^2 = 542.28$ ($df = 425$), $p < .01$, Bollen-Stine- $p = .26$, $\chi^2/df = 1.28$, CFI = .92, TLI = .90, SRMR = .0549, RMSEA = .034, CI = [.024;.042], $p_{close} = 1.00$). Lediglich der Faktor *Lernen lernen* hat einen signifikanten Pfadkoeffizienten, die Varianzaufklärung des Regressionsmodells beträgt 18.2 %.

Tabelle 105: Regression Wiederholen auf EPI

Skala	partielles Regressionsgewicht		SE	CR	P
	nicht standardisiert	standardisiert			
RN	.17	.14	.16	1,07	.28
AW	.09	.08	.12	.72	.47
S/OW	.08	.07	.13	.65	.51
SK	.08	.08	.14	.53	.60
VK	.10	.06	.18	.56	.58
LL	.64	.30	.28	2.30	.02

5.3.2.5 Planen

Für die Regression der Strategie *Planen* liegt die Modellpassung im guten bis akzeptablen Bereich ($\chi^2 = 426.78$ ($df = 366$), $p = .02$, Bollen-Stine- $p = .56$, $\chi^2/df = 1.17$, CFI = .95, TLI = .94, SRMR = .0521, RMSEA = .026, CI = [.012;.036], $p_{close} = 1.00$). Allerdings ist keiner der Pfadkoeffizienten signifikant. Lediglich der Pfadkoeffizient von *Lernen lernen* liegt

nahe an der Signifikanzgrenze, allerdings ist die Höhe des standardisierten Pfadkoeffizienten nicht über der gewählten Grenze von .30. Die Varianzaufklärung beträgt 14.6 %.

Tabelle 106: Regression Planen auf EPI

Skala	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
RN	.09	.07	.17	.52	.60
AW	.15	.14	.13	1.13	.26
S/OW	.13	.12	.14	.93	.36
SK	.07	.08	.15	.50	.62
VK	-.12	-.08	.20	-.61	.54
LL	.51	.25	.27	1.89	.06

5.3.2.6 Überwachen

Die Modellpassung liegt im guten bis akzeptablen Bereich ($\chi^2 = 408.97$ ($df = 338$), $p = .005$, Bollen-Stine- $p = .44$, $\chi^2/df = 1.21$, CFI = .94, TLI = .93, SRMR = .0520, RMSEA = .029, CI = [.017;.039], $p_{close} = 1.00$). Dieses Modell zeigt drei signifikanten Pfadkoeffizienten, und zwar die Faktoren *Lernen lernen*, *Reflexive Natur des Wissens* und *Soziale Komponente des Wissens*. Das Modell hat eine Varianzaufklärung von 42.5 %.

Tabelle 107: Regression Überwachen auf EPI

Skala	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
RN	.49	.49	.17	2.88	***
AW	.08	.09	.11	.71	.48
S/OW	-.19	-.22	.12	-1.63	.10
SK	.34	.46	.14	2.34	.02
VK	-.30	-.23	.18	-1.61	.11
LL	.56	.34	.25	2.22	.03

Anmerkung: ***: $p < .001$

5.3.2.7 Regulieren

Auch hier liegt die Modellanpassung im guten bis akzeptablen Bereich ($\chi^2 = 460.64$ ($df = 366$), $p = .001$, Bollen-Stine- $p = .32$, $\chi^2/df = 1.26$, CFI = .92, TLI = .90, SRMR = .0527, RMSEA = .033, CI = [.022;.041], $p_{close} = 1.00$). Dieses Modell zeigt zwei signifikante Pfadkoeffizienten, dies sind die der Faktoren *Lernen lernen* und *Reflexive Natur des Wissens*. Der Pfadkoeffizient von *Absolutes Wissen* auf *Regulieren* liegt auf der Signifikanzgrenze, wobei aber die Höhe des Pfadkoeffizienten weit unterhalb der Schwelle von .30 liegt. Die Varianzaufklärung ist 39.4 %.

Tabelle 108: Regression Regulieren auf EPI

Skala	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
RN	.34	.33	.13	2.55	.01
AW	.21	.22	.11	1.95	.05
S/OW	-.05	-.05	.11	-.45	.65
SK	-.12	-.15	.12	-1,03	.30
VK	-.17	-.13	.16	-1,11	.27
LL	.45	.26	.21	2.14	.03

5.3.2.8 Anstrengung

Die Fitindizes zeigen eine gute bis akzeptable Modellanpassung ($\chi^2 = 600.48$ ($df = 456$), $p < .001$, Bollen-Stine- $p = .24$, $\chi^2/df = 1.32$, CFI = .90, TLI = .89, SRMR = .0554, RMSEA = .036, CI = [.028;.044], $p_{close} = .99$). Lediglich der Faktor *Reflexive Natur des Wissens* hat einen signifikanten Effekt auf die Strategie *Anstrengung*, wobei der Pfadkoeffizient die gesetzte Schwelle von .30 genau erreicht. Die Varianzaufklärung beträgt 22.8 %.

Tabelle 109: Regression Anstrengung auf EPI

Skala	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
RN	.34	.30	.15	2.32	.02
AW	.19	.18	.11	1.66	.10
S/OW	-.14	-.13	.12	-1.18	.24
SK	-.18	-.20	.13	-1.34	.18
VK	.08	.06	.15	.53	.60
LL	.28	.15	.21	1.36	.17

5.3.2.9 Aufmerksamkeit

Auch hier liegt der Fit des Modells im guten bis akzeptablen Bereich ($\chi^2 = 515.90$ ($df = 425$), $p = .01$, Bollen-Stine- $p = .48$, $\chi^2/df = 1.21$, CFI = .95, TLI = .95, SRMR = .0538, RMSEA = .030, CI = [.019;.038], $p_{close} = 1.00$). Allerdings hatte keiner der epistemologischen Überzeugungen einen signifikanten Einfluss auf die Strategie *Aufmerksamkeit*. Die Varianzaufklärung beträgt 3 %.

Tabelle 110: Regression Aufmerksamkeit auf EPI

Skala	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
RN	-.19	-.12	.20	-.97	.33
AW	-.18	-.13	.15	-1.18	.24
S/OW	.17	.12	.16	1.07	.29
SK	.09	.08	.17	.55	.58
VK	.10	.05	.22	.43	.67
LL	.04	.02	.27	.15	.88

5.3.2.10 Zeitmanagement

Die Modellpassung liegt im guten bis akzeptablen Bereich ($\chi^2 = 434.05$ ($df = 366$), $p = .008$, Bollen-Stine- $p = .49$, $\chi^2/df = 1.19$, CFI = .95, TLI = .95, SRMR = .0514, RMSEA = .028, CI = [.015;.037], $p_{close} = 1.00$). Lediglich der Faktor *Soziale Komponente des Wissens* zeigte einen signifikanten Pfadkoeffizienten, wobei dieser genau die Signifikanzgrenz erreicht. Die Varianzaufklärung des Modells beträgt 12.3 %.

Tabelle 111: Regression Zeitmanagement auf EPI

Skala	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
RN	.25	.15	.22	1.15	.25
AW	.14	.09	.18	.80	.42
S/OW	.06	.04	.18	.32	.75
SK	.40	.32	.20	2.00	.05
VK	-.48	-.20	.31	-1.54	.12
LL	.57	.21	.33	1.74	.08

5.3.2.11 Lernumgebung

Die Modellpassung liegt im guten bis akzeptablen Bereich ($\chi^2 = 495.80$ ($df = 395$), $p < .001$, Bollen-Stine- $p = .32$, $\chi^2/df = 1.26$, CFI = .93, TLI = .91, SRMR = .0529, RMSEA = .032, CI = [.022;.041], $p_{close} = 1.00$). Nur die Faktoren *Lernen lernen* und *Reflexive Natur des Wissens* wiesen signifikante Pfadkoeffizienten auf, wobei dessen Größe beim Faktor *Lernen lernen* genau auf der Grenze von .30 lag, die Varianzaufklärung beträgt 29.1 %.

Tabelle 112: Regression Lernumgebung auf EPI

Skala	partielles Regressionsgewicht		SE	CR	p
	nicht standardisiert	standardisiert			
RN	.34	.34	.13	2.59	.01
AW	.13	.14	.10	1.33	.19
S/OW	-.02	-.02	.10	-.17	.86
SK	.11	.15	.11	1.03	.30
VK	-.13	-.10	.15	-.86	.39
LL	.50	.30	.22	2.29	.02

5.3.2.12 Literatur

Die Fitindizes des Modells liegen bis auf den SRMR alle im Bereich guter Modellpassung ($\chi^2 = 413.20$ ($df = 366$), $p = .05$, Bollen-Stine- $p = .643$, $\chi^2/df = 1.13$, CFI = .96, TLI = .95, SRMR = .0520, RMSEA = .023, CI = [.004;.034], $p_{close} = 1.00$). Hier zeigten lediglich die Faktoren *Reflexive Natur des Wissens* und *Lernen lernen* einen signifikanten Pfadkoeffizienten, wobei beim Faktor *Lernen lernen* der standardisierte Regressionskoeffizient nicht über .30 liegt. Die Varianzaufklärung des Regressionsmodells beträgt 21.9 %.

Tabelle 113: Regression Literatur auf EPI

Skala	partielles Regressionsgewicht		SE	CR	p
	Nicht standardisiert	standardisiert			
RN	.63	.38	.22	2.82	.01
AW	.09	.06	.16	.53	.60
S/OW	-.28	-.19	.17	-1.60	.11
SK	.28	.23	.19	1.50	.13
VK	-.36	-.16	.27	-1.36	.18
LL	.62	.22	.33	1.86	.06

5.3.2.13 Zusammenfassung der Regressionsmodelle

In Tabelle 114 und Tabelle 115 sind die Ergebnisse für die kognitiven, metakognitiven und ressourcenbezogenen Lernstrategien zusammengefasst. In den Zeilen der Tabellen finden sich die epistemologischen Überzeugungen und in den Spalten die Lernstrategien. Die signifikanten Pfadkoeffizienten in den jeweiligen Regressionsmodellen sind fett hervorgehoben, kursiv bedeutet, dass die Signifikanzgrenze nur knapp verfehlt wurde, alle anderen Regressionskoeffizienten sind normal gedruckt.

In Tabelle 116 und Tabelle 117 sind die Regressionskoeffizienten ihrer Effektstärke entsprechend im Sinn der Klassifikation von Cohen (1988) aufgeführt. Analog zu Tabelle 114 und

Tabelle 115 sind diejenigen Regressionskoeffizienten fett gedruckt, die signifikant waren und diejenigen kursiv gedruckt, die die Signifikanzgrenze nur knapp verfehlten.

Tabelle 114: Zusammenfassung der Regressionsergebnisse für die kognitiven und metakognitiven Lernstrategien

Skala	Kognitive Lernstrategien				Metakognitive Lernstrategien		
	Elaborieren	Kritisches Prüfen	Organisieren	Wiederholen	Planung	Überwachen	Regulieren
RN	.44	.42	.08	.14	.07	.49	.34
AW	.15	-.02	.13	.08	.14	.09	.21
S/OW	-.15	.01	-.12	.07	.18	-.22	-.05
SK	.04	.22	.07	.08	.08	.46	-.15
VK	-.03	-.19	-.08	.06	-.08	-.23	-.13
LL	.01	.04	.35	.30	.25	.34	.26
R ²	16.1 %	15.2 %	15.6 %	18.2 %	14.6 %	42.5 %	39.4 %

Tabelle 115: Zusammenfassung der Regressionsergebnisse für die ressourcenbezogenen Lernstrategien

Skala	interne Ressourcen			externe Ressourcen	
	Anstrengung	Aufmerksamkeit	Zeitmanagement	Lernumgebung	Literatur
RN	.34	-.12	.15	.34	.38
AW	.19	-.13	.09	.14	.06
S/OW	-.14	.12	.04	-.02	-.19
SK	-.18	.08	.32	.15	.23
VK	.08	.05	-.20	-.10	-.16
LL	.15	.02	.21	.30	.22
R ²	22.8 %	3.0 %	12.3 %	29.1 %	21.9 %

Tabelle 116: Klassifikation der Effektstärken nach Cohen (1988) für kognitive und metakognitive Lernstrategien.

Skala	Kognitive Lernstrategien				Metakognitive Lernstrategien		
	Elaborieren	Kritisches Prüfen	Organisieren	Wiederholen	Planung	Überwachen	Regulieren
RN	mittel	mittel	klein	mittel	klein	mittel	mittel
AW	mittel	klein	klein	klein	mittel	klein	<i>mittel</i>
S/OW	mittel	klein	klein	klein	mittel	mittel	klein
SK	klein	mittel	klein	klein	klein	mittel	mittel
VK	klein	mittel	klein	klein	klein	mittel	mittel
LL	klein	klein	mittel	mittel	<i>mittel</i>	mittel	mittel
R ²	mittel	mittel	mittel	mittel	mittel	groß	groß

Tabelle 117: Klassifikation der Effektstärken nach Cohen (1988) für die ressourcenbezogenen Lernstrategien

Skala	interne Ressourcen			externe Ressourcen	
	Anstrengung	Aufmerksamkeit	Zeitmanagement	Lernumgebung	Literatur
RN	mittel	mittel	mittel	mittel	mittel
AW	mittel	mittel	klein	mittel	klein
S/OW	mittel	mittel	klein	klein	mittel
SK	mittel	klein	mittel	mittel	mittel
VK	klein	klein	mittel	klein	mittel
LL	mittel	klein	mittel	mittel	mittel
R ²	mittel	klein	mittel	groß	mittel

5.4 Diskussion der Regressionsanalysen

Bevor die Ergebnisse der Regressionsanalysen analysiert werden, soll eine kurze Diskussion der Untersuchung der Skalen mithilfe der konfirmatorischen Faktorenanalysen vorangehen. Die differenzierte Betrachtung der Ergebnisse der CFA der LIST-Skalen ist insofern von besonderer Bedeutung, da bis jetzt keine Prüfung der faktoriellen Validität dieser Skalen mithilfe der CFA-Methodik existiert und die Qualität der LIST-Skalen einen wesentlichen Stützpfeiler zur Untersuchung der prädikativen Validität epistemologischer Überzeugungen darstellt. Zur Sicherung der faktoriellen Validität wurden die einzelnen LIST-Skalen jeweils einer konfirmatorischen Faktorenanalyse unterzogen. Dabei sind zwei Besonderheiten zu beachten, welche die Skala zur Erfassung der metakognitiven Lernstrategien und die Skala zur Erfassung des Lernens mit Studienkollegen betreffen.

Zunächst musste die Struktur der Skala zur Erfassung der metakognitiven Lernstrategien geklärt werden. Diese soll die drei Aspekte *Planen*, *Überwachen* und *Regulieren* von Lernvorgängen erfassen. In der ursprünglichen Analyse von Wild und Schiefele (1994) wurden diese drei Faktoren nicht gefunden und entsprechend wurden die Items zur Erfassung dieser Strategie in einer gemeinsamen Skala *Metakognitive Strategien* zusammengefasst. Boerner et al. (2005) konnten hingegen die dreidimensionale Struktur der Skala *Metakognitive Strategien* nachweisen. Aus diesem Grund wurde hier evaluiert, ob sich ebenfalls eine dreidimensionale Struktur finden lässt. Eine EFA ergab eine klare 3-faktorielle Ladungsmatrix, mit der Besonderheit, dass Item 61, welches ursprünglich zur Erfassung der Strategie *Überwachen* gehört, zusammen mit den Items der Strategie *Regulieren* auf einem Faktor lud. Die Skala *Überwachen* beinhaltet Items, welche die Intention ausdrücken, das Verstehen von Lernmaterial zu überwachen, während die Skala *Regulieren* Items beinhaltet, welche die Aufdeckung von Verständnisschwierigkeiten und Wissenslücken sowie das Ergreifen entsprechender Maßnahmen beinhalten. Insbesondere das Item 13 bezieht sich ebenso wie das Item 61 auf Wissenslücken. Aus diesen Erwägungen heraus kann Item 61 im Sinne einer Regulation des zu rekapitulierenden Stoffes verstanden werden, allerdings ohne entsprechende Maßnahmen zur Regulation dieser Wissenslücken zu erwähnen. Deswegen wurde Item 61 im Folgenden zu der Strategie *Regulieren* hinzugefügt. Ein entsprechendes 3-faktorielles CFA-Modell führte zu guten bis akzeptablen Fitindizes.

Es zeigten sich auch Probleme mit der Skala *Lernen mit Studienkollegen*. Diese wies eine extrem schlechte Modellpassung auf, was möglicherweise auf die Itemformulierung zurück-

zuführen war. Da die Prüfung dieser Möglichkeit im Rahmen der vorliegenden Studie nicht möglich war, wurde auf eine weiterführende Verwendung in den Regressionsanalysen epistemologischer Überzeugungen auf die erhobenen Lernstrategien verzichtet.

Überdies wurden nur geringfügige Änderungen an den Skalen des LIST vorgenommen, die auf den Ergebnissen der konfirmatorischen Faktorenanalyse beruhen. Die vorgenommenen Modifikationen waren einheitlich die Eliminierung bestimmter Items, wenn die Residuen bzw. Modifikationsindizes Hinweise auf korrelierende Messfehler gaben. Um ein einheitliches Vorgehen mit den vorhergehenden Untersuchungen des EBQ- bzw. EPI-Fragebogens zu gewährleisten, wurden entsprechende Items aus den LIST-Skalen entfernt. Alternativ wäre es auch möglich gewesen, die gefundenen korrelierenden Messfehler in dem entsprechenden Regressionsmodell explizit zu berücksichtigen (Kennedy, 2007).

Allerdings muss auf eine Einschränkung der faktoriellen Validierung des LIST hingewiesen werden, die überdies auch für die im Folgenden diskutierten Regressionsmodelle gilt. Die Modifikationen der LIST-Skalen wurden nur an einer Stichprobe durchgeführt und nicht mithilfe einer weiteren unabhängigen Stichprobe validiert. Die aufgeführten Modifikationen der LIST-Skalen sowie die Regressionsmodelle sind somit nur im Kontext der vorliegenden Stichprobe gültig, eine Replikation der Ergebnisse ist folglich notwendig. Da die Modifikationen an den LIST-Skalen aber nur gering sind (in keiner Skala wurde mehr als ein Item entfernt) und da bei den Skalen zur Erfassung der metakognitiven Lernstrategien theoretische Annahmen für das Vorliegen einer dreidimensionalen Struktur sprechen, kann das Fehlen einer separaten Validierung der modifizierten LIST-Skalen als vertretbar angesehen werden.

Die mithilfe von Strukturgleichungen modellierten Regressionsanalysen hatten fast alle gute bis akzeptable Modellpassungen. Die einzige Ausnahme hiervon waren der TLI-Wert der Regression der EPI-Faktoren auf die Strategien *Kritisches Prüfen* und *Anstrengung*, der Wert des TLI lag bei nur .89 und somit unter der Grenze von .90 für eine akzeptable Modellpassung. Da alle anderen Fitindizes aber im Bereich guter bis akzeptabler Modellpassung lagen, die Unterschreitung marginal war und der TLI wegen seiner fehlenden Normierung nur informell zur Beurteilung der Modellpassung betrachtet wird, wurde das Modell beibehalten.

Auffällig in allen Regressionsmodellen sind jedoch die Abweichungen der Wahrscheinlichkeit des χ^2 -Wertes von der mithilfe der Bollen-Stine-Bootstrap-Methode korrigierten Wahr-

scheinlichkeit. Da der Bollen-Stine-Bootstrap die Effekte von Abweichungen der manifesten Variablen von der Normalverteilung korrigiert, sind offensichtlich zu beachtende Abweichungen der Daten von der Normalverteilungsannahme vorhanden. Abweichungen von der Normalverteilung führen dazu, dass die Schätzung der Standardfehler der Modellparameter fälschlicherweise zu klein ausfällt, sodass bei der Signifikanztestung diese eher signifikant werden, was bei der Interpretation der standardisierten Pfadkoeffizienten beachtet werden muss. Diese Parameter können somit überhöht sein und einen zu starken Zusammenhang epistemologischer Überzeugungen und Lernstrategien angeben.

Zur Prüfung der Konstruktvalidität wurde eine Regressionsanalyse durchgeführt, die als Strukturgleichungsmodell formuliert wurde. Wird die Konstruktvalidität geprüft, sollte eine Versuchsplanung erfolgen und unter Berücksichtigung der α - und β -Fehler die optimale Stichprobengröße für den erwünschten Effekt bestimmt werden (Hartig et al., 2008). Ein solches Vorgehen war im Rahmen der durchgeführten Untersuchung nicht möglich, da die Daten in Verbindung mit einer anderen Untersuchung erhoben wurden, die ein solches Vorgehen ausschloss. Daher wurde zur Konstruktvalidierung eine abgewandelte Strategie gewählt. Einerseits sollten die Regressionskoeffizienten mindestens mittlere Effektstärken im Sinne der Klassifikation von Cohen (1988) erreichen und andererseits auch signifikant sein. Damit sollte verhindert werden, dass signifikante, aber von der Effektstärke nur kleine Effekte in die Konstruktvalidierung eingehen. Jeweils abwägend beurteilt werden sollten diejenigen Regressionskoeffizienten, die knapp die Signifikanzgrenze und/oder knapp die Effektstärke verfehlten.

Die Ergebnisse der Regressionsanalyse sprechen nicht für einen generellen Zusammenhang epistemologischer Überzeugungen mit Lernstrategien. Von 72 möglichen Regressionskoeffizienten sind nur 14 signifikant, zwei Pfadkoeffizienten verfehlten nur knapp die Signifikanzgrenze von .05, diese waren die Pfadkoeffizienten von *Lernen lernen* auf die Strategien *Planen* und *Literatur*. Am häufigsten zeigten die beiden Faktoren *Reflexive Natur des Wissens* sowie *Lernen lernen* einen signifikanten Pfad zu den jeweiligen Lernstrategien, sonst zeigte nur der Faktor *Soziale Komponente des Wissens* bei der Regression auf die Lernstrategie *Zeitmanagement* einen signifikanten Regressionskoeffizienten. Der Faktor *Absolutheit des Wissens* verfehlte bei der Regression der Lernstrategie *Regulieren* nur sehr knapp die Signifikanzgrenze. Die Ergebnisse sprechen also eher für einen Zusammenhang nur einiger epistemologischer Überzeugungen mit Lernstrategien. In Tabelle 116 und Tabelle 117 wurden die

Regressionskoeffizienten entsprechend ihrer Effektstärke aufgeführt. In diese Betrachtung wurden auch diejenigen Regressionskoeffizienten einbezogen, welche nur knapp die Signifikanzgrenze verfehlten. Wie die beiden Tabellen verdeutlichen, liegen nur kleine bis mittlere Effektstärken vor. Die signifikanten Regressionskoeffizienten zeigen auch gleichzeitig mittlere Effektstärken.

Allerdings sollte neben den kleinen bis mittleren Effektstärken noch die Varianzaufklärung beachtet werden, die bei Strukturgleichungsmodellen durch die quadrierte multiple Korrelation angegeben wird (s. Tabelle 114 und Tabelle 115). Die Varianzaufklärung ist in fast allen Modellen gering, lediglich in der Regression epistemologischer Überzeugungen auf die beiden metakognitiven Strategien *Überwachen* und *Regulieren* und die Strategie *Lernumgebung* finden sich nennenswerte Varianzaufklärungen. Der Varianzaufklärung lässt sich entsprechend des multiplen Regressionskoeffizienten R bzw. dessen quadrierten Wert R^2 anhand der Cohen'schen Klassifikation einteilen (s. Tabelle 116 und 116). Lediglich bei drei von 12 der zuvor genannten Regressionsmodelle wurden große Effektstärken gefunden. Alle übrigen Varianzaufklärungen lagen im Bereich mittlerer Effektstärken, im Fall der Regression auf die Lernstrategie *Aufmerksamkeit* war die Effektstärke überdies so klein, dass kein Zusammenhang zwischen epistemologischen Überzeugungen und dieser Lernstrategie angenommen werden kann. Dies deutet darauf hin, dass epistemologische Überzeugungen im Allgemeinen wenig dazu geeignet sind, Lernstrategien vorherzusagen.

Die Resultate sind im Einklang mit den bisherigen Befunden über den Zusammenhang von epistemologischen Überzeugungen und Lernstrategien (vgl. Kapitel 1.6.2). So fand Cano (2005) Zusammenhänge, deren Pfadkoeffizienten betragsmäßig zwischen .07 und .30 lagen. In dem Pfadmodell von Phan (2006) fanden sich Zusammenhänge in den betragsmäßigen Größen von .16 bis .29, wobei sich in dieser Studie zeigte, dass nur zwei der vier durch den EBQ erfassten epistemologischen Überzeugungen signifikante Pfadkoeffizienten zu den Lernstrategien aufweisen. Auch die Varianzaufklärungen sind für die beiden erhobenen Lernstrategien ähnlich gering wie in den eigenen Befunden, sie betragen 10 % bzw. 14 %. Interessant bei der Studie von Phan (2006) allerdings ist, dass epistemologische Überzeugungen starke Zusammenhänge mit Faktoren des kritischen Denkens aufweisen. Diese liegen im Bereich zwischen .15 und .61, und auch die Varianzaufklärungen betragen zwischen 24 % und 47 %. Bei Phan (2008) finden sich wieder schwache Zusammenhänge von epistemologischen

Überzeugungen mit Strategien selbstregulierten Lernens, die Varianzaufklärung war auch in dieser Studie mit 13.2 % gering.

Von besonderer Bedeutung für den Vergleich mit den hier durchgeführten Untersuchungen sind die Studien von Urhahne und Hopf (2004) sowie die Studie von Dahl et al. (2005). In diesen beiden Studien wurden die Lernstrategien mithilfe des LIST bzw. des MSLQ, aus dem der LIST entwickelt wurde, erhoben. In der Studie von Urhahne und Hopf (2004) konnten ausgehend vom Fragebogen von Conley et al. (2004), der auf dem Modell von Hofer und Pintrich (1997) beruht, sowie dem LIST-Fragebogen von Wild und Schiefele (1994) Zusammenhänge zwischen Elaborations- und Kontrollstrategien mit dem Faktor *Rechtfertigung* nachgewiesen werden. Die beiden bivariaten Korrelationen betragen .34 für Elaborationsstrategien und .35 für Kontrollstrategien, entsprechend liegen die Effektstärken dieser beiden bivariaten Korrelationen im mittleren Bereich und die aufgeklärte Varianz beträgt jeweils 11.6 % bzw. 12.3 %. Die restlichen Korrelationen befinden sich im kleinen Bereich. Die Befunde von Urhahne und Hopf konvergieren insoweit mit den hier ermittelten Ergebnissen, als dass nur kleine bis mittlere Effektstärken vorkamen und nur eine Dimension epistemologischer Überzeugungen (*Rechtfertigung*) einen stärkeren Zusammenhang mit Elaborations- und Kontrollstrategien aufwies. Das Resultat spiegelt wider, dass die Dimension *Reflexive Natur des Wissens* ein signifikanter Prädiktor mit mittlerer Effektstärke für die Elaborationsstrategien ist. Vergleichbar mit den Befunden der eigenen Untersuchung ist auch der von Urhahne und Hopf gefundene Zusammenhang von *Rechtfertigung* mit Kontrollstrategien. Die Kontrollstrategien sind die Zusammenfassung der in der vorliegenden Studie getrennten metakognitiven Lernstrategien *Planung*, *Überwachen* und *Regulieren*. Dies ähnelt dem Ergebnis, wonach *Reflexive Natur des Wissens*, *Soziale Komponente des Wissens* und *Lernen lernen* signifikante Prädiktoren für die metakognitiven Strategien *Überwachen* und *Regulieren* sind.

Die Untersuchungsergebnisse von Dahl et al. (2005) sind ebenfalls konform mit den erzielten Ergebnissen, betroffen hiervon sind die Varianzaufklärungen der Regressionsmodelle. In der Studie von Dahl et al. (2005) wurden aus dem EBQ vom Schommer abgeleitete Skalen als Prädiktoren zur Vorhersage der mithilfe des MSLQ von Pintrich et al. (1991) erfassten Lernstrategien untersucht. Als epistemologische Überzeugungen wurden die vier Faktoren *Simple Knowledge*, *Certain Knowledge*, *Quick Learning* und *Fixed Ability* identifiziert. Diese Faktoren gingen in Regressionsmodelle als Prädiktoren ein, die jeweiligen Lernstrategien des MSLQ stellten die Kriterien dar. Die aufgeklärten Varianzen dieser Regressionsmodelle lagen

zwischen 11 % und 25 %. Ebenfalls zu finden ist, dass es nur einige epistemologische Überzeugungen sind, die signifikante Prädiktoren darstellten. Bei Dahl et al. (2005) waren dies *Simple Knowledge* und *Fixed Ability*, d. h. auch in dieser Studie wird das Befundmuster repliziert, dass nur einige epistemologische Überzeugungen zur Vorhersage von Lernstrategien dienen, diese Vorhersage aber relativ wenig Varianzaufklärung hat. Allerdings werden die Ergebnisse durch die Verwendung des EBQ-Fragebogens von Schommer eingeschränkt, da die faktorielle Struktur dieses Fragebogens unklar ist (vgl. Kapitel 1.3.1, Clarebout et al., 2001, DeBacker et al., 2008) und ebenso wird die Vergleichbarkeit durch die Verwendung von herkömmlichen Regressionsanalysen in Gegensatz zu den hier eingesetzten Strukturgleichungsmodellen eingeschränkt.

Eine Einschränkung, die sich aber nicht nur auf die vorliegende Untersuchung bezieht, liegt in der Art der Erfassung von Lernstrategien. Die Lernstrategien wurden durch Selbstreporte in Verbindung mit einer fünfstufigen Ratingskala erfasst. Hierbei mussten die Probanden auf der Ratingskala angeben, wie häufig eine bestimmte Lerntätigkeit ausgeführt wird. Dies stellt aber keine objektive Erfassung dar, sondern gibt lediglich die von den Versuchspersonen subjektiv berichteten Häufigkeiten der eingesetzten Lernstrategien an. Artelt (2000) konnte zeigen, dass die von Probanden selbst berichteten Lernstrategien nicht mit den gezeigten Lernstrategien übereinstimmen. Die Erfassung der Lernstrategien durch Selbstberichte ist verschiedenen Einflüssen ausgesetzt, die eine Interpretation des Zusammenhangs von epistemologischen Überzeugungen und Lernstrategien erschweren. Diese Einflüsse lassen sich in drei Arten einteilen.

Die erste Art von Einflüssen geht die auf die Interpretation der Items und der Ratingskala durch den Probanden zurück. Sie beschreiben z. B. auf die Interpretation der Bedeutung der Ratingskala-Stufen durch die Versuchspersonen, oder auf Erinnerungseffekte, wie oft das betreffende Verhalten in der Vergangenheit durchgeführt wurde oder aber auf welchen Zeitraum sich die anzugebenden Häufigkeiten beziehen. Ebenso können unterschiedliche Assoziationen mit den Begriffen, die in einem Item vorkommen, zu einer uneinheitlichen Interpretation der Items bei mehreren Versuchspersonen führen (vgl. Fisseni, 2004; Krosnick, 1999).

Die zweite Art von Einflüssen stellen die aus der Fragebogenmethode bekannten Antworttendenzen dar, wie z. B. die Tendenz zur Mitte und die Ja- bzw. Nein-Sage-Tendenz. Die Tendenz zur Mitte tritt vor allem dann auf, wenn das zu beurteilende Merkmal dem Probanden

wenig vertraut ist und dieser deswegen keine differenzierte Einstufung auf der Ratingskala treffen kann. Die Ja- bzw. Nein-Sage-Tendenz hängt in der Regel mit anderen Persönlichkeitseigenschaften der Probanden zusammen, die dazu führen, dass der Proband eher generell zustimmt bzw. nicht zustimmt. Die genannten Tendenzen können aber auch dann auftreten, wenn sich die Probanden in ihren Urteilen unsicher sind. Zusätzlich kann die Reihenfolge der Items zu Primacy-Recency-Effekten führen, die ebenfalls zu Urteilsverzerrungen führen können.

Die letzte Art von Einflüssen stellen Urteilsverzerrungen dar, die durch Testverfälschung entstehen. Hierzu gehören die Selbstdarstellung und die soziale Erwünschtheit. Selbstdarstellung entsteht dadurch, dass die Bearbeitung eines Fragebogens durch die Versuchsperson als kommunikativer Akt erlebt wird, in dem die Versuchsperson anderen Personen etwas über sich mitteilen kann und sich somit die Chance zum Impression-Management bietet. Beim Impression-Management wählt die Versuchsperson bewusst aus, welche Information sie preisgibt und in welcher Art die Präsentation dieser Informationen erfolgt. Auf diese Weise versucht die Versuchsperson, einen von ihr gewünschten Eindruck zu erzielen (Mummendey, 1990). Soziale Erwünschtheit bezeichnet die Tendenz, gemäß den in der Gesellschaft erwünschten Normen zu antworten, um so sozialer Verurteilung zu entgehen (vgl. Bortz & Döring, 2006, Mummendey, 1999).

Die Einschätzung der Lernstrategien durch die Versuchsperson selbst stellt also keine numerisch korrekte Messung dar, vielmehr sollten die Einschätzungen der Probanden auf einer Ratingskala als eine grobe Einschätzung der selbstwahrgenommenen Ausprägung des eingeschätzten Merkmals angesehen werden. Der Sachverhalt wird auch noch dadurch verkompliziert, dass auch die Selbstbeurteilung epistemologischer Überzeugungen nicht zwingend ein genaues Bild der epistemologischen Überzeugungen liefert. Ähnlich wie bei Lernstrategien ebenfalls gibt es Befunde, die nahelegen, dass sich selbstberichtete epistemologische Überzeugungen von nicht mit den epistemologischen Überzeugungen übereinstimmen, die sich im Handeln von Personen ausdrücken (Leach et al., 2004, Limón, 2006b). Der Zusammenhang von epistemologischen Überzeugungen und Lernstrategien sollte, wenn überhaupt, somit in erster Linie tendenziell und nicht ausschließlich über die Höhe der Regressionskoeffizienten beurteilt werden⁵².

⁵² An dieser Stelle lässt sich die Frage stellen, ob die üblicherweise angewandten Methoden wie Regressionsanalysen oder Strukturgleichungsmodell adäquate Methoden darstellen, da diese mindestens eine metrische Skalendignität verlangen. An einer metrischen Skalendignität lassen aber die im Text aufgeführten Urteilsverzerrungen,

Somit kann in Bezug auf die letzte Forschungsfrage zum EPI festgehalten werden, dass in erster Linie die Faktoren *Lernen lernen* und *Reflexive Natur des Wissens* Konstruktvalidität im Sinne konkurrenter Validität zeigen. Hinweise auf Konstruktvalidität liefern darüber hinaus die Faktoren *Soziale Komponente des Wissens* und *Absolutheit des Wissens*.

Diese Art der vorgenommenen Konstruktvalidierung durch Bestimmung der konkurrenten Validität weist aber auch einige Schwächen auf. Zum einen muss auf die tentative Natur der Skalen zur Erfassung epistemologischer Überzeugungen hingewiesen werden, die auf jeden Fall einer weiteren Revision bedürfen, wie die Untersuchungen in Kapitel 5.1 verdeutlicht haben. Zum anderen wurden keine expliziten Hypothesen aufgestellt, wie dies für eine exakte Konstruktvalidierung erforderlich gewesen wäre (vgl. Hartig et al., 2008). Die hier aufgestellte Hypothese des Zusammenhangs zwischen epistemologischen Überzeugungen und Lernstrategien ist in erster Linie von dem Wirkmodell von Hofer (2001) abgeleitet und daher auch von vergleichbarer allgemeiner Natur wie dieses Modell, in dem nur ein genereller Zusammenhang zwischen epistemologischen Überzeugungen und Lernstrategien postuliert wird. Diese Ergebnisse geben somit nur an, dass zwischen einigen epistemologischen Überzeugungen und einigen Lernstrategien ein Zusammenhang besteht. Diese Allgemeinheit der Hypothesen hat ihren Grund in der Neuartigkeit des EPI und des dahinter stehenden Modells. So war es zum einen unklar, ob sich die Faktorenstruktur des Fragebogens empirisch bestätigen ließ, sodass eine a-priori-Ableitung von Hypothesen über den Zusammenhang der EPI-Faktoren mit Lernstrategien nicht möglich war. Zum anderen liegen bis jetzt keine empirischen Befunde über den Zusammenhang der EPI-Faktoren mit anderen Konstrukten vor, insbesondere keine Untersuchungen über den Zusammenhang dieses Fragebogens mit Lernstrategien, wie sie mithilfe des LIST erfasst werden. Alle Zusammenhangshypothesen haben somit letztendlich explorativen Charakter. Der Forderung nach einer exakten Hypothesentestung sollte somit in weiteren Untersuchungen zur Validität des EPI nachgekommen werden. Die vorliegende Untersuchung kann als Ansatzpunkt zur weitergehenden Validierung des von Moschner und Gruber (im Druck) entwickelten EPI- bzw. FEE-Fragebogens genutzt werden, um gerichtete Zusammenhangshypothesen mit anderen Konstrukten aufzustellen und zu testen.

Gegen die Art der hier durchgeführten Validierung wurde auch grundsätzliche Kritik vorgebracht (Borsboom et al., 2004, Borsboom et al., 2009a). Danach ist die Korrelation eines

die bei der Beantwortung eines Items auftreten können, zweifeln. Insofern ist auch eine Reflexion der eingesetzten statistischen Methoden angebracht.

Konstrukts mit anderen Konstrukten kein Indikator für die Validität eines Instruments. Vielmehr muss zur Sicherstellung der Validität eines Instruments gewährleistet werden, dass dieses auch diejenige Eigenschaft misst, die es zu messen intendiert. Ein Konstrukt ist somit valide, wenn zwischen diesem Konstrukt und den Messwerten, die das Konstrukt erfassen sollen, eine kausale Beziehung besteht. Um nun die Validität der Messung sicherzustellen, ist eine theoretische Erklärung der Prozesse notwendig, die zu den Messergebnissen führen. Benötigt wird also eine Theorie, wie der Messvorgang stattfindet. Ein solches Vorgehen, das eine theoretische Fundierung des Zusammenhangs zwischen Konstrukt und Messung vorgibt, wurde auch von Embretson (1983) vorgeschlagen. Nach Borsboom et al. (2004) ist die theoretische Bestimmung des Messprozesses aber ein psychologisches und kein methodologisches Unterfangen. Für die Erfassung epistemologischer Überzeugungen bedeutet dieses, dass eine Theorie benötigt wird, die den Prozess beschreibt, welcher der Itembeantwortung zugrunde liegt. Möglichkeiten zur Untersuchung der Itembeantwortung bieten z. B. die Methoden, die in der Fragebogenforschung bei kognitiven Interviews zum Einsatz kommen (Sudman, Bradburn & Schwarz, 1996, Tourangeau, Rips & Rasinski, 2000).

Unabhängig von der vorgebrachten Kritik an den herkömmlichen Verfahren zu Validierung eines Messinstruments sprechen die vorgestellten Zusammenhänge der EPI-Faktoren in jedem Fall für die Relevanz wenigstens einiger epistemologischer Überzeugungen im Lernprozess und damit auch für eine pädagogische Relevanz wenigstens einiger epistemologischer Überzeugungen. Vor allem die Zusammenhänge der Faktoren *Reflexive Natur des Wissens* und *Lernen lernen* lassen sich für die Praxis des Lernens verwerten. Wird die Überzeugung von Lernenden gestärkt, dass entsprechende Lerntechniken erworben werden können, erhöht dies auch die Wahrscheinlichkeit, dass entsprechende Lerntechniken angewandt werden. Wird die Überzeugung des Lernenden verstärkt, dass sich Wissen durch neue Erfahrungen und durch eine kritische Auseinandersetzung damit weiterentwickelt, erhöht dies die Wahrscheinlichkeit, dass die zu lernenden Inhalte besser elaboriert und auf logische Konsistenz überprüft werden. Gleichzeitig sollte sich die Wahrscheinlichkeit des metakognitiven Monitorings der eigenen Lernleistung erhöhen.

6. Allgemeine Diskussion

In den vorhergehenden Kapiteln wurden die Forschungsfragen zu EBQ- und EPI-Fragebogen beantwortet. Dabei wurde eine Reihe weiterer Fragen aufgeworfen, speziell in Bezug auf die theoretische Abgrenzung des Konstruktes der epistemologischen Überzeugung von anderen Konstrukten, die Problematik der Erfassung epistemologischer Überzeugungen sowie die Frage nach der Struktur epistemologischer Überzeugungen. In den folgenden Kapiteln werden die bei der Diskussion der einzelnen Ergebnisse aufgeworfenen Fragen unter allgemeinen Gesichtspunkten diskutiert, die über die Forschungsfragen hinausgehen.

Bei der weiterführenden Diskussion wird in der folgenden Reihenfolge vorgegangen: Zuerst wird die Abgrenzung von epistemologischen Überzeugungen zu anderen Konstrukten vorgenommen. Hierbei wird ein Kriterium vorgestellt, das eine solche Abgrenzung erlaubt. Anschließend wird für die in den Fragebogen genannten Dimensionen mithilfe dieses Kriteriums geprüft, ob es sich um epistemologische Überzeugungen handelt oder nicht. Die bei der Analyse der Fragebogen aufgetretenen Probleme werden im darauf folgenden Kapitel besprochen, wobei sowohl die Problematik der Itemgestaltung als auch die Frage nach der Validität im Zentrum steht. Diese beiden Aspekte sind nicht nur bei der Erfassung epistemologischer Überzeugungen relevant, sondern betreffen die gesamte Fragebogenforschung in der Psychologie.

Der nächste Punkt, der besprochen wird, handelt von der normativen Sicht epistemologischer Überzeugungen, welche der Methodik der Erfassung epistemologischer Überzeugungen mittels Fragebogens implizit zugrunde liegt. Die normative Sichtweise wird am Beispiel der Überzeugungen zur Sicherheit des Wissens diskutiert. Hierbei werden die Begriffe Produktivität und Korrektheit eingeführt und es wird ausgehend von diesen beiden Begriffen argumentiert, dass epistemologische Überzeugungen dann fortgeschritten sind, wenn sie eine der jeweiligen Situation angepasste Beurteilung von Aussagen mit sich bringen. Der nächste Kapitel beschäftigt sich mit anderen Aspekten der normativen Vorstellungen, nämlich den sozialen Aspekten epistemologischer Überzeugungen. Während üblicherweise angenommen wird, dass Personen mit fortgeschrittenen epistemologischen Überzeugungen der Meinung sind, dass Wissen am besten selbst erworben wird, wird hier argumentiert, dass die Verteilung von Wissen in einer Gesellschaft zu der Unmöglichkeit führt, alles Wissen durch eigene Erfahrung zu erwerben. Epistemologische Überzeugungen sind aus diesem Grund notwendige Mechanismen zur Beurteilung der Aussagen und der Glaubwürdigkeit verschiedener Wissens-

quellen. Ein weiterer Aspekt der normativen Sichtweise, der mit dem Wissenserwerb durch eigene Erfahrung in Zusammenhang steht, ist die Referenz auf externe Quellen. In der normativen Sichtweise wird angenommen, dass ein solcher Bezug auf eine externe Quelle keine fortgeschrittene epistemologische Überzeugung ist. Im Zusammenhang mit der im vorhergehenden Kapitel vorgetragenen Argumentation wird hier gezeigt, dass der Bezug auf externe Wissensquelle ein unverzichtbarer Bestandteil von epistemologischen Überzeugungen ist und nicht als epistemologisch naiv gelten kann.

Im nächsten Kapitel wird die Problematik besprochen, dass die mittels Fragebogen erhobenen epistemologischen Überzeugungen nicht mit denen übereinstimmen müssen, die z. B. durch Beobachtungen erschlossen wurden. Die Diskussion dieser Frage führt schließlich zu einem Ansatz epistemologischer Überzeugungen, der mit den vorher besprochenen Überlegungen kongruent ist. Epistemologische Überzeugungen werden hier als Schemata konzipiert.

Bei dem eigenen Ansatz handelt es sich dabei nicht um eine strukturelle Theorie, welche spezielle Dimensionen epistemologischer Überzeugungen postuliert. Aus den diskutierten Aspekten wird eine kognitive Architektur epistemologischer Überzeugungen abgeleitet und ebenso werden geeignete methodische Instrumente zur Untersuchung vorgestellt. Das methodische Instrumentarium, das hier vorgestellt wird, ist ein Latent-State-Trait-Ansatz epistemologischer Überzeugungen. Der eigene Ansatz ist dabei nicht vollständig, es handelt sich vielmehr um eine Skizze, die eine Möglichkeit aufzeigt, in welcher Richtung in Zukunft die kognitiven Aspekte epistemologischer Überzeugungen untersucht werden und welche methodischen Instrumente angewandt werden können, was zu einer partiellen Überschneidung der bei den einzelnen Punkten vorgetragenen Überlegungen führen kann, sodass verschiedene Befunde und Überlegungen an mehreren Stellen jeweils in einem unterschiedlichen Zusammenhang vorkommen können. Letztendlich werden in den beiden letzten Kapiteln die Argumentationsstränge zu dem eigenen Ansatz epistemologischer Überzeugungen zusammengeführt.

6.1 Abgrenzung epistemologischer Überzeugung

Zunächst wird die konzeptionelle Abgrenzung epistemologischer Überzeugungen gegenüber anderen Konstrukten diskutiert. Diese Abgrenzung wird an erster Stelle thematisiert, da sie Auswirkungen auf die anschließend zu besprechende Gestaltung von Items zur Messung epistemologischer Überzeugungen hat. Die Diskussion, aus welchen Komponenten sich epistemologische Überzeugungen zusammensetzen, betrifft insbesondere die Frage, ob Überzeugungen

zum Lernen bzw. Überzeugungen zur Intelligenz Bestandteil epistemologischer Überzeugungen sind oder nicht (vgl. Kapitel 1.3.1). Nach Schommer-Aikins (2004) sind Überzeugungen zum Lernen und zur Begabung Bestandteil epistemologischer Überzeugungen, da die Beurteilung von Autoritäten als Quelle von Wissen die Beurteilung der Begabung dieser Autoritäten erfordert. Ebenso müssen die eigenen Fähigkeiten zum Lernen im Vergleich zu der Fähigkeit der Autoritäten eines Fachs beurteilt werden. Nach Moschner und Gruber (2003) bestehen epistemologische Überzeugungen aus den Komponenten Begabung, Wissen und Lernen, die in den Alltagsvorstellungen von Lernenden und Lehrenden verbunden werden. Gruber und Moschner (2003) zufolge üben die epistemologischen Überzeugungen von Lernenden Einfluss auf den Lernerfolg aus und die epistemologischen Überzeugungen von Lehrenden wirken auf die Gestaltung von Lernangeboten, die Unterstützung bei der Selektion geeigneter Lernstrategien von Lernenden sowie die Förderung von Begabung und Wissenserwerb. Dabei spielt aber auch der jeweilige fachwissenschaftliche Erkenntnisstand eine Rolle. Moschner et al. (2005) fassen in ihrer Konzeption epistemologische Überzeugungen sehr weit und integrieren in ihren Fragebogen auch Überzeugungen zur Kultur- und Geschlechtsabhängigkeit von Wissen, zu Wertigkeit und Nutzen von Wissen und zur sozialen Natur von Wissen. Wie in Kapitel 5.2 dargelegt wurde, bilden insbesondere Überzeugungen zur Kultur- und Geschlechtsabhängigkeit von Wissen keine genuinen epistemologischen Überzeugungen.

Rozendaal et al. (2001), Greene, Azevedo und Torney-Purta (2008) sowie Chinn, Buckland und Amarapungavan (2011) nutzen zur Abgrenzung des Konstrukts ein Kriterium, das auf der philosophischen Disziplin der Epistemologie gründet. Steup (2008) charakterisiert Epistemologie als Fach, das sich mit den Bedingungen und Quellen von Wissen auseinandersetzt. Die Epistemologie untersucht, welche Bedingungen erfüllt sein müssen bzw. welche Quellen gegeben sein müssen, damit eine Aussage als gerechtfertigt angenommen werden kann. Eine gerechtfertigte Aussage wird in der Epistemologie dann als Wissen aufgefasst. Formal lässt sich eine gerechtfertigte Aussage wie folgt darstellen (Herzog, 2012): Eine Person S weiß, dass p genau dann, wenn (1) p wahr ist, (2) S überzeugt ist, dass p der Fall ist, und (3) S in ihrer Überzeugung gerechtfertigt ist. Dieses Abgrenzungskriterium muss nun auf die psychologische Forschung zu epistemologischen Überzeugungen übertragen werden. Hierzu kann auf den Begriff der propositionalen Einstellung zurückgegriffen werden (Gardenne, 2004, Schwitzgebel, 2008).

Propositionale Einstellungen sind Aussagen einer Person über sich oder die Außenwelt, zu der diese in irgendeiner Art und Weise Stellung bezieht. Propositionale Einstellungen bestehen aus einem Modus sowie einem mental repräsentierten Inhalt. Propositionale Einstellungen lassen sich in der Form „Person [Verb], dass ...“ ausdrücken, wobei im ersten Teilsatz der Modus steht, im zweiten Teilsatz der mental repräsentierte Inhalt (Westermann, 2000, Gardenne, 2004). Z. B. wird durch den Satz „Der Schüler Hans glaubt, dass die Aussage des Lehrers wahr ist“ eine propositionale Einstellung ausgedrückt. Solch eine propositionale Einstellung wird als gerechtfertigt bezeichnet, wenn diese sich entweder durch internale oder externale Rechtfertigung als wahr bezeichnen lässt. Eine internale Rechtfertigung ist dabei durch Reflexion zu erkennen und eine externale Rechtfertigung beruht auf direkter Wahrnehmung des betreffenden Sachverhaltes. So ist z. B. die propositionale Einstellung „Ich weiß, dass der Himmel blau ist“, durch eine direkte Wahrnehmung des blauen Himmels zu rechtfertigen. Es wäre aber auch eine internale Rechtfertigung möglich, wenn die entsprechenden Gesetzmäßigkeiten aus der Physik herangezogen werden, welche die blaue Farbe des Himmels erklären.

Weiterhin relevant ist der Begriff der Überzeugung. Eine Überzeugung ist die Bewertung eines Objekts oder eines Sachverhalts (Fiske & Taylor, 1991). Der zu bewertende Sachverhalt wird durch eine propositionale Einstellung zusammen mit deren Rechtfertigungskriterien gebildet, z. B. wird durch die propositionale Einstellung „Ich weiß, dass der Himmel blau ist“, zusammen mit den physikalischen Gesetzmäßigkeiten, welche diese Aussage rechtfertigen, ein Sachverhalt gebildet. Rozendaal et al. (2001) argumentieren nun, dass sich eine Überzeugung dann als epistemologisch bezeichnen lässt, wenn sich diese auf die Beurteilung der Rechtfertigung von Aussagen bezieht. Somit ist ein Kriterium gegeben, das es erlaubt, epistemologische Überzeugungen von anderen Überzeugungen und Konstrukten zu unterscheiden.⁵³ Dieses Kriterium wurde auch von Richter (2007) und Bromme und Stahl (2008) vorgeschlagen. Ebenso plädieren Greene, Azevedo und Torney-Purta (2008) sowie Chinn, Buckland und Amarapungavan (2011) für eine Abgrenzung auf Basis des aus der philosophischen Epistemologie übernommen Rechtfertigungskriteriums. Ausgehend von diesem Kriterium bestimmen Rozendaal et al. (2001) für das Modell von Hofer und Pintrich (1997), ob die darin enthaltenen Dimensionen epistemologische Überzeugungen sind oder nicht.

⁵³ Im weiteren Verlauf des Textes wird zur Vereinfachung der Formulierung nur noch von der Beurteilung von Aussagen gesprochen. Gemeint ist damit aber immer, dass sich die Beurteilung auf die Rechtfertigung einer Aussage bezieht.

Tabelle 118: Dimensionen nach Hofer und Pintrich (1997) Erfüllung des Rechtfertigungskriteriums

Dimension	Natur des Wissen (<i>nature of knowledge</i>)		Prozess des Wissenserwerbs (<i>nature of knowing</i>)	
	Certainty of knowledge	Simplicity of knowledge	Source of know- ledge	Justification of knowing
Erfüllung des Kriteriums	Kriterium erfüllt	Kriterium nicht erfüllt	Kriterium erfüllt	Kriterium erfüllt

Für die Dimension *Certainty of Knowledge* ist das Kriterium erfüllt, da sich diese Dimension darauf bezieht, ob Wissen eher vorläufig und sich entwickelnd oder sicher ist. Der Bezug zur Rechtfertigung ergibt sich nach Rozendaal et al. (2001) aus der Beurteilung einer Behauptung als sicher oder unsicher, welche die Existenz einer objektiven Wahrheit voraussetzt, die internal oder external gerechtfertigt werden kann. Die Dimension *Simplicity of Knowledge* zielt auf die Frage, ob Wissen aus vernetzten oder einzelnen Aussagen besteht. Dies hat offensichtlich keinen Bezug zur Rechtfertigung einer Aussage. *Source of Knowledge* beschreibt, ob eine Person annimmt, dass Wissen von außen durch eine Autorität vermittelt oder durch den Wissenden konstruiert wird. Hier wird auf die zuvor genannten Unterschiede von internaler oder externaler Rechtfertigung in der philosophischen Epistemologie Bezug genommen. Die Dimension zielt also auf die Frage ab, ob Wissen außerhalb der Person existiert und erkennbar ist (externale Rechtfertigung) oder ob Wissen durch die Person selbst konstruiert wird (internale Rechtfertigung). Sie bezieht sich somit auf die Rechtfertigung von Aussagen und kann als epistemologische Überzeugung behandelt werden. *Justification of knowledge* bezieht sich darauf, ob zur Prüfung einer Behauptung empirische Evidenz oder die Aussagen von Autoritäten benutzt werden, wodurch sie auch einen unmittelbaren Bezug zur Rechtfertigung einer Aussage hat.

Dieses Kriterium gestattet es aber nicht nur, bestimmte Konstrukte als epistemologische Überzeugung zu charakterisieren, es bietet auch die Möglichkeit, einzelne Fragebogenitems auf ihre Relevanz zu überprüfen. Hierbei würde für jedes Item geprüft werden, ob sich die in dem Item beschriebene Handlung bzw. Überzeugung auf die Rechtfertigung von Aussagen bezieht. Ist dies nicht der Fall, so ist das Item nicht zur Erfassung epistemologischer Überzeugungen geeignet. Somit ist dieses Kriterium nicht nur in der Abgrenzung epistemologischer Überzeugungen gegenüber anderen Konstrukten dienlich, sondern es gibt auch ein nützliches Instrument bei der Entwicklung von Items für die Erfassung epistemologischer Überzeugungen an die Hand.

Allerdings setzt die Rechtfertigung einer Aussage ein bestimmtes, themenbezogenes Vorwissen voraus (Bromme et al., 2008). Dies geht auch mit der Rolle von Vorwissen konform, die in der Wissenspsychologie im Rahmen des Wissenserwerbs diskutiert wird (vgl. Krause & Stark, 2006). Nach Bromme et al. (2008) ist vor allem das Wissen über Methoden, die angewandt wurden, um zu einer Aussage zu gelangen, sowie Wissen über die Art und Weise, wie bestimmte Aussagen verbreitet werden (z. B. in welchen Medium), für die Rechtfertigung relevant. Neben dem themenbezogenen Vorwissen spielen ontologische Annahmen über das betreffende Fachgebiet eine Rolle. Ontologie ist die philosophische Disziplin, die sich mit dem Seienden befasst, ontologische Annahmen beziehen sich demnach auf bestimmte Kategorien, nach welchen die Realität einer Disziplin gegliedert wird (Bromme et al., 2008). Ontologische Annahmen sind wichtig, um die Angemessenheit von Aussagen in einem Themengebiet zu beurteilen. So stellen Motivation und Volition geeignete Erklärungen in einem psychologischen Kontext dar, die aber nicht im Kontext physikalischer Erklärungen oder bei evolutionstheoretischen Erklärungen angebracht sind. Epistemologische Überzeugungen und die damit verbundene Rechtfertigung von Aussagen sind somit immer mit dem Vorwissen einer Person und deren ontologischen Annahmen über eine Disziplin verknüpft.

Im Folgenden wird dieses Rechtfertigungskriterium genutzt und die in den vorangegangenen Kapiteln untersuchten Dimensionen des EBQ und EPI daraufhin überprüft, ob diese als epistemologische Überzeugung gewertet werden können, wobei sich die folgende Betrachtung auf die ursprüngliche Form der Fragebögen bezieht und nicht auf die Dimensionen, die sich aus den empirischen Analysen ergeben haben. Eine Betrachtung der Items des EBQ, deren Wortlaut in Anhang A zu finden ist, zeigt, dass sich lediglich drei Items auf das Rechtfertigungskriterium beziehen. Das Item „Man kann fast alles glauben, was man für das Studium liest“, bezieht sich auf die Wissensquelle, wobei als Argument zur Rechtfertigung sich hier auf publizierte Aussagen in der einschlägigen Literatur bezogen wird. Das Item lässt aber offen, welche Studien zur Rechtfertigung der Aussagen herangezogen werden sollen. Das Item „Wissenschaftler können letztendlich zur Wahrheit kommen“ bezieht sich auf eine Autorität, die zur Rechtfertigung einer Aussage herangezogen wird. Allerdings enthält das Item keinen Hinweis, wer als Autorität angesehen werden kann. Das Item „Manchmal muss man die Antworten des Dozenten akzeptieren, auch wenn man sie nicht versteht“, bezieht sich ebenfalls auf die Rechtfertigung durch eine Autorität, wobei dieses Item allerdings noch Aspekte des eigenen Verständnisses einer Aussage beinhaltet.

Bei der Analyse der Items des EPI-Fragebogens, deren Wortlaut in Anhang B zu finden ist, stellt sich die Situation ähnlich dar wie bei den Items des EBQ. Zudem sind auch die ersten beiden beim EBQ genannten Items im EPI enthalten. Die Items „Wenn Probleme besonders strittig sind, wendet man sich am besten an Experten“ und „Die Antworten von Dozenten müssen als wahr akzeptiert werden“ beziehen sich ebenfalls auf Autoritäten zur Rechtfertigung von Aussagen. Das Item „Wenn man etwas in einem Lehrbuch liest, dann kann man auch sicher sein, dass es wahr ist“, bezieht sich auf die Literatur zur Rechtfertigung von Aussagen. Das Item „Nur veröffentlichte Erkenntnisse können als Wissen gelten“ bezieht sich ebenfalls auf die wissenschaftliche Literatur, wobei hier nur solche Aussagen als gültig angesehen werden, die in wissenschaftlicher Fachliteratur veröffentlicht wurden. Das Item lässt es allerdings offen, welche Kriterien, wie z. B. peer review, an diese Art von Literatur angelegt werden, damit diese eine geeignete Wissensquelle darstellt.

Neben einer Betrachtung des Rechtfertigungskriteriums auf Itemebene besteht noch die Möglichkeit, die theoretischen Konstrukte daraufhin zu überprüfen, ob deren Definition das Rechtfertigungskriterium erfüllt und ob diese überhaupt eine epistemologische Überzeugung darstellen, was exemplarisch für den revidierten EPI-Fragebogen vorgenommen wird (vgl. Kapitel 5.1.3 und 5.1.4). Der revidierte EPI-Fragebogen besteht aus den in Tabelle 119 aufgeführten Faktoren, die eine Übersicht über die Dimensionen des revidierten EPI-Fragebogens und deren Definition in Anlehnung an Moschner et al. (2005) sowie eine Übersicht über die Erfüllung des Rechtfertigungskriteriums für jede Dimension enthält.

Reflexive Natur des Wissens gibt die Veränderlichkeit von Wissen in Bezug auf die Erfahrung des Lernenden wieder. Diese Dimension kann als epistemologische Überzeugung angesehen werden, da die Beurteilung von neuem Wissen sowie die Veränderung von bestehendem Wissen es erfordert, die Rechtfertigung neu erworbener bzw. vorhandener Wissensinhalte zu prüfen. Ebenso kann es erforderlich sein, bestehendes oder neu erworbenes Wissen zu revidieren. Diese Dimension kann somit als eine interne Prüfung der Rechtfertigung angesehen werden.

Absolutheit von Wissen bezieht sich auf die Existenz von Wahrheit bzw. auf die Möglichkeit, wahre Aussagen zu erhalten. Dieses Konstrukt zielt also unmittelbar auf das Erkennen des Wahrheitsgehaltes von Aussagen ab und ist damit direkt mit dem Rechtfertigungskriterium im Sinne einer externalen Rechtfertigung verbunden. Dieses Konstrukt beschreibt somit eine epistemologische Überzeugung.

Tabelle 119: Dimensionen des revidierten EPI-Fragebogens sowie die Erfüllung des Rechtfertigungskriteriums

Dimension	Charakterisierung	Kriterium erfüllt/nicht erfüllt
Reflexive Natur des Wissens	Darunter wird die Veränderlichkeit von Wissen in Abhängigkeit von der Perspektive und Erfahrung des Lernenden verstanden	Kriterium erfüllt
Absolutheit des Wissens	Bezieht sich auf die Existenz von Wahrheit bzw. auf die Möglichkeit, wahre Aussagen zu erhalten	Kriterium erfüllt
Sicherheit/Objektivität des Wissens	Beschreibt das Ausmaß, mit dem Wissen als sicher und wenig veränderlich sowie als objektiv angesehen wird	Kriterium erfüllt
Soziale Komponente des Wissens	Ausmaß der sozialen Komponente des Wissens	Kriterium erfüllt
Value of knowledge	Konstrukt welches auf die Fragestellung abzielt, inwieweit Individuen die Bedeutung von Wissen an dessen gesellschaftlichem Nutzen messen	Kriterium nicht erfüllt
Geschlechtsspezifische Wissenszugänge	Einstellungen der Versuchspersonen zum Zusammenhang von Geschlecht und Wissen bzw. Wissenserwerb	Kriterium nicht erfüllt
Kulturspezifische Wissenszugänge	Einfluss, den die Kultur auf die Überzeugungen über den Prozess des Wissenserwerbs und über das Wissen hat	Kriterium nicht erfüllt
Lernen lernen	Einfluss von Lerntechniken auf den Erfolg beim Wissenserwerb	Kriterium nicht erfüllt

Das Konstrukt *Sicherheit/Objektivität des Wissens*⁵⁴ bezeichnet das Ausmaß, mit dem Wissen als sicher und wenig veränderlich sowie als objektiv angesehen wird, analog der Dimension *Certainty of Knowledge* von Hofer und Pintrich (1997). Nach Rozendaal et al. (2001) ist für *Certainty of Knowledge* der Bezug zum Rechtfertigungskriterium gegeben, da die Beurteilung einer Behauptung als sicher oder unsicher die Existenz einer objektiven Wahrheit voraussetzt, die internal oder external gerechtfertigt werden kann. Bei *Sicherheit/Objektivität des Wissens* handelt es sich somit um eine epistemologische Überzeugung.

⁵⁴ Die Dimension *Sicherheit/Objektivität des Wissens* ist im ursprünglichen EPI-Fragebogen nicht enthalten (vgl. Anhang B), diese wurde im Rahmen der Untersuchungen in Kapitel 5.1.3 eingeführt.

An dieser Stelle wird zudem deutlich, dass *Absolutheit von Wissen* und *Sicherheit/Objektivität des Wissens* von ihrer theoretischen Konzeption her gesehen, nicht unabhängig voneinander sind. Dies geht auch mit den empirischen Befunden aus Kapitel 5.1 einher, wonach die beiden Faktoren zu $.32$ korrelierten. Dies zeigt, dass die Sicherheit bzw. Objektivität nicht unabhängig von der Absolutheit des Wissens beurteilt werden kann. Die Entscheidung, ob eine Aussage als sicher oder unsicher und objektiv angesehen werden kann, setzt voraus, dass der Wahrheitsgehalt einer Aussage beurteilt werden kann.

Die *Soziale Komponente des Wissens* beschreibt das Ausmaß, inwieweit Wissen als gültig angesehen wird, das mit anderen geteilt wird. Dies kann bedeuten, dass beispielsweise die Publikation von Wissen in der Literatur oder die Anerkennung bestimmter Aussagen durch die Scientific Community zur Rechtfertigung herangezogen wird. Dieser soziale Aspekt der Rechtfertigung wird auch von Rozendaal et al. (2001) betont, die darauf hinweisen, dass Wissen nicht in einem sozialen Vakuum entsteht. Vielmehr entsteht Wissen auch durch einen sozialen Prozess, an dem mehrere Individuen beteiligt sind und der zu einem Konsens über die Rechtfertigung einer Aussage führt. Auf die Wichtigkeit solcher Prozesse in modernen Gesellschaften, in denen Wissen in hohem Maße zwischen den Mitgliedern dieser Gesellschaft verteilt ist, weisen auch Bromme et al. (2008) hin. Danach kann das heute verfügbare, umfangreiche Spezialwissen von einer Person nicht mehr in dem Ausmaß erworben werden, um internal bewertet zu werden (für eine ausführliche Diskussion der sozialen Aspekte epistemologischer Überzeugungen im Sinn verteilten Wissens siehe Kapitel 6.3.4). Demnach ist man bei der Rechtfertigung von Aussagen auf zugängliche Publikationen zum Thema von anerkannten Experten des betreffenden Faches, Lehrbücher etc. angewiesen. Die Beurteilung sozialer Elemente des Wissens wird somit zu einer Notwendigkeit bei der Rechtfertigung von Aussagen und neu erlernten Wissensinhalten. Das Konstrukt *Soziale Komponente des Wissens* ist daher eine epistemologische Überzeugung.

Value of Knowledge, also die Beurteilung von Wissen als für eine Gesellschaft nutzbringend oder nicht, stellt keine epistemologische Überzeugung dar. Der gesellschaftliche Nutzen weist keinen Bezug zur Rechtfertigung von Aussagen auf.

Für die beiden Konstrukte *Geschlechtsspezifische Wissenszugänge* und *Kulturspezifische Wissenszugänge* wurde schon bei der Analyse des EPI-Fragebogens gezeigt, dass diese keine epistemologischen Überzeugungen darstellen. Auch in Bezug auf das Rechtfertigungskriteri-

um stellen sie keine epistemologischen Überzeugungen dar. Allerdings stellt sich für diese Konstrukte die Frage, inwieweit epistemologische Überzeugungen von Geschlecht bzw. Kultur abhängig sind.

Lernen lernen, das den Einfluss von Lerntechniken auf den Erfolg beim Wissenserwerb thematisiert, stellt im Sinne des Rechtfertigungskriteriums ebenfalls keine epistemologische Überzeugung dar. Aber auch für dieses Konstrukt ist die Beziehung mit den epistemologischen Überzeugungen interessant.

Auf der Basis des Rechtfertigungskriteriums lässt sich somit bestimmen, welches Konstrukt als epistemologische Überzeugung gelten kann. Insbesondere lässt sich die Aussage von Schommer-Aikins (2004) bewerten, wonach Überzeugungen zum Lernen und zur Begabung Bestandteil epistemologischer Überzeugungen sind. Auch wenn die Beurteilung von Autoritäten als Quelle von Wissen sowie die Beurteilung der Begabung dieser Autoritäten für die Beurteilung der Rechtfertigung einer Aussage eine Rolle spielen kann, muss dies nicht zwangsläufig der Fall sein. Zu fragen ist hier, in welcher Beziehung die Einstellung über eine bestimmte Autorität einer Wissenschaft sowie die Beurteilung der Begabung dieser Autorität mit der eigentlichen Beurteilung einer wissenschaftlichen Aussage im Zusammenhang stehen. In der Sichtweise von Moschner et al. (2005), die das Konstrukt *Lernen lernen* unter der Perspektive der Beeinflussbarkeit der eigenen Lernfähigkeit sehen, ist zu fragen, inwieweit solche Einstellungen im Zusammenhang mit epistemologischen Überzeugungen stehen. Z. B. könnte hier von Interesse sein, ob die Einstellung, dass bestimmte Lerntechniken erworben werden können, im Zusammenhang mit der internalen oder externalen Rechtfertigung einer Aussage stehen.

Bis hier wurde die Abgrenzung epistemologischer Überzeugungen von anderen Konstrukten beschrieben. Diese Abgrenzung dient in erster Linie dazu, konzeptuelle Klarheit in die Diskussion zu bringen, aus welchen Konstrukten epistemologische Überzeugungen bestehen. Diese konzeptionelle Klarheit ist überdies wichtig bei der Konstruktion von Items zur Erfassung epistemologischer Überzeugungen.

6.2 Erfassung epistemologischer Überzeugungen

6.2.1 Itemgestaltung

Bei den Auswertungen des EBQ sowie des EPI sind zum Teil massive Probleme in Bezug auf die Items aufgetreten. Beim EBQ-Fragebogen führte dies unter anderem zu einem Ausschluss von Items vor Durchführung der explorativen Faktorenanalyse. Bei der Analyse des EPI-Fragebogens führten nicht auf das Konstrukt bezogene Items häufig zu nicht passenden Messmodellen bzw. führte die Itemformulierung, z. B. bei semantisch ähnlichen Items, zu korrelierenden Messfehlern. Im Folgenden wird die Konstruktion von Items zur Messung epistemologischer Überzeugungen näher betrachtet. Dabei wird auch ein wichtiger Punkt zur Beurteilung der Validität von Skalen zur Erfassung epistemologischer Überzeugungen besprochen.

Eine der auffälligsten Beobachtungen während der faktorenanalytischen Untersuchung des EPI-Fragebogens war der durch die Formulierung der Items hervorgerufene Methodenfaktor bei dem Faktor *Geschlechtsspezifische Wissenszugänge*. Dieser wurde durch die auf Unterschiedlichkeit und Verschiedenheit der Geschlechter lautende Formulierung der Items bedingt. Dieses Beispiel verdeutlicht, inwieweit die gegensätzliche Formulierung von Items die dimensionale Struktur eines Messmodells bedingen kann (vgl. Brown, 2006, Hazlett-Stevens, Ulman & Craske, 2004, DiStefano & Motl, 2009). Daher sollten positiv und negativ formulierte Items vermieden werden, um ggf. auftretende Methodeneffekte zu vermeiden. Überdies kann unter Umständen die unterschiedliche Polung der Items von Versuchspersonen nicht richtig erkannt werden. Dies führt dazu, dass die Versuchspersonen nicht in der intendierten Richtung antworten werden. Dies kann zu einer Ablehnung eines an sich passenden Messmodells führen (Woods, 2006).

Ein anderes – bei der Analyse sowohl des EPI als auch des EBQ auftretendes – Phänomen waren korrelierende Messfehler zwischen zwei Items. Korrelierende Messfehler deuten auf gemeinsame Varianz zwischen zwei oder mehreren Items hin, die neben dem eigentlichen Einfluss der latenten Variablen existiert. Korrelierte Messfehler stellen daher einen Methodenfaktor dar, welcher neben der latenten Variablen die Kovariation der Items bedingt. Die Gründe für korrelierende Messfehler sind vielfältig. Z. B. können die Zusammenfassung von Items zu Subgruppen oder Reihenfolgeeffekte innerhalb eines Fragebogens oder ein bestimmtes Antwortverhalten aufgrund gleichartiger Formulierungen in einer Teilmenge der

Items oder Gemeinsamkeiten wegen kurzzeitiger Einstellungen der Befragten gegenüber der zu messenden Eigenschaft, die nur bei einigen Items auftritt, bei anderen aber nicht, zu korrelierenden Messfehlern führen (Green & Yang, 2009). Diesen Sachverhalten muss bei der Entwicklung von Skalen Rechnung getragen werden. So muss beispielsweise auf gleichlautende Formulierungen geachtet werden und es sollten auch keine Items verwendet werden, die sich inhaltlich überlappen. Reihenfolgeeffekte können mithilfe entsprechender Designs untersucht werden. Ein weiteres Problem bei der Itemformulierung betrifft die Verwendung von Superlativen. Solche Items kommen z. B. im EPI-Fragebogen vor. Beispiele hierfür sind die Items „Es gibt unumstößliche Wahrheiten“ und „Es gibt wissenschaftliche Erkenntnisse, die immer gültig sein werden“. Diese können zur Verwendung extremer Antwortkategorien führen, was eine Varianzeinschränkung bedingen kann, welche die Eigenschaften der Messmodelle negativ beeinflussen kann (Nye, Newman & Joseph, 2010).

Für die Formulierung von Items hilfreich ist eine Methodik, welche die systematische Generierung von Items ermöglicht. Einen Ansatzpunkt bietet hierzu das Vorgehen von Gerber (2004), welche die Konstruktion eines Fragebogens zur Erfassung epistemologischer Überzeugungen mithilfe der Facettentheorie (Borg, 1992) vorgenommen hat. Bei diesem Vorgehen wird das interessierende Konstrukt zuerst theoretisch abgegrenzt und anschließend werden facettentheoretische Abbildungssätze generiert, die zur Formulierung der Items dienen (s. Kapitel 1.3.2 für das Beispiel eines Abbildungssatzes). Die Abgrenzung kann mittels des in Kapitel 6.1 genannten Rechtfertigungskriteriums erfolgen. Auf der Basis dieser Abbildungssätze werden anschließend Items entwickelt. Dieses Vorgehen bietet den Vorteil, dass Items systematisch entwickelt werden und damit einhergehend gleichartige Formulierungen und inhaltliche Überlappungen vermieden werden, die zu Methodenfaktoren führen können.

Die Formulierung von Abbildungssätzen bringt einen zusätzlichen Vorteil. Ein Abbildungssatz baut auf den Facetten auf. Eine Facette stellt eine Regel dar, die angibt, in welcher Art und Weise ein interessierender Bereich in verschiedene Klassen aufgeteilt werden kann (Bortz & Döring, 2006). Z. B. wird in dem Abbildungssatz zur Dimension *Sicherheit des Wissens* in Kapitel 1.3.2 die Beschaffenheit eines Wissensbereichs als relativ oder sicher facettiert. Eine solche Facettierung gestattet es aber auch, die zwei gegensätzlichen Pole einer bestimmten Dimension epistemologischer Überzeugungen darzustellen. Im genannten Beispiel würde die Dimension Sicherheit des Wissens von den beiden Polen relativ und sicher aufgespannt werden.

Eine Facettierung ermöglicht somit auch die Bildung bipolarer Ratingskalen, sodass auch eine Erfassung epistemologischer Überzeugungen mithilfe eines semantischen Differenzials (Osgood, Suci & Tannenbaum, 1957, Snider & Osgood, 1969) erfolgen kann, durch das die konnotativen Aspekte eines Beurteilungsobjektes erfasst werden sollen. Nach Snider und Osgood (1969) bezieht sich die denotative Bedeutung auf die funktionalen Aspekte eines Begriffs und die konnotative Bedeutung auf die assoziativen und evaluativen Aspekte eines Begriffs. Diese Art, die Methode des semantischen Differenzials zu nutzen, ist der Linguistik entnommen und dient dazu, die mit einem bestimmten Begriff assoziierten Begriffe sowie deren Evaluation zu ermitteln, um diese im semantischen Raum zu verorten (Crystal, 1993). Hierzu wird Probanden ein Begriff vorgegeben, die diesen dann wiederum mithilfe bipolarer Ratingskalen beurteilen müssen, wobei die jeweilige Bewertung als Pole der bipolaren Skala vorgegeben ist. Ein semantisches Differenzial zur Erfassung epistemologischer Überzeugungen wurde von Stahl und Bromme (2007) im CAEB-Fragebogen umgesetzt. Stahl und Bromme (2007) bauen dabei auf der Unterscheidung denotativer und konnotativer Wortbedeutungen auf. Im Hinblick auf epistemologische Überzeugungen bezieht sich die denotative Bedeutung auf das explizite Wissen einer Person über Wissenschafts- und Erkenntnistheorie, während die konnotative Bedeutung sich auf die persönliche Beurteilung des Wissens bezieht. Stahl und Bromme (2007) nehmen an, dass eine Person beim Lesen eines Items dieses auf seine konnotative Bedeutung bezieht. Betrachtet man epistemologische Überzeugungen in Anlehnung an Hofer und Pintrich (1997) als Persönliche Theorien über Wissen, sind es genau die konnotativen Aspekte, die mithilfe eines Instruments erfasst werden sollen. Nach Stahl und Bromme (2007) in Verbindung mit Snider und Osgood (1969) können konnotative Aspekte epistemologischer Überzeugungen also durch ein semantisches Differenzial erfasst werden.

Die Erfassung epistemologischer Überzeugungen mithilfe eines semantischen Differenzials bietet den Vorteil, dass neben den konnotativen Aspekten auch denotative Aspekte epistemologischer Überzeugungen erfasst werden können, da sich diese nicht klar trennen lassen (Stahl & Bromme, 2007). Allerdings ist es mittels eines semantischen Differenzials aber auch nicht möglich, die reine Denotation eines Begriffs zu ermitteln (Crystal, 1993). Das Antwortformat mithilfe von bipolaren Adjektiven hilft zudem, Messfehler zu vermeiden, die auf den Eigenschaften der Itemformulierung basieren. Ein Nachteil des semantischen Differenzials ist, dass die Beurteilung auf einer bipolaren Ratingskala den meisten Versuchspersonen wenig vertraut ist, sodass die verlangte Beurteilung Schwierigkeiten bereiten kann. Diesem Problem kann aber mittels einer geeigneten Instruktion vorgebeugt werden (Bortz & Döring, 2006). Ein wei-

terer Nachteil ist, dass der semantischen Raum, in dem ein Begriff verortet werden soll, immer nur aus den Dimensionen bestehen kann, für die entsprechende bipolare Begriffspaare bei der Konstruktion bedacht wurden. Somit werden die möglichen Bedeutungsdimensionen durch die Person festgelegt, welche das semantische Differenzial erstellt. Letztendlich ist die Verortung eines Begriffs im semantischen Raum entscheidend von der Vorstellung der Person abhängig, die das semantische Differenzial erstellt hat. Werden dann konnotative Bedeutungsdimensionen auf faktorenanalytischem Weg abgeleitet, wie dies beim CAEB geschehen ist, bedeutet dies, dass die gefundenen Dimensionen von den Items abhängig sind, welche die Faktorenanalyse beinhaltet. Vereinfacht ausgedrückt besagt dies, dass eine Faktorenanalyse nur solche Dimensionen liefern kann, für die von Anfang an Begriffe zur Messung ausgewählt wurden. Diese Auswahl hängt aber von der theoretischen Position des jeweiligen Forschers ab. Im wissenschaftstheoretischen Sinn ist diese Auswahl somit theoriegetränkt (Westermann, 2000). Dieser Nachteil ist aber nicht nur auf das semantische Differenzial bzw. die Erfassung epistemologischer Überzeugungen bezogen, sondern betrifft die gesamte psychologische Forschung, in der diese Methode Anwendung findet.

6.2.2 Validitätsaspekte

Neben einer Itemformulierung, die es erlaubt, das Konstrukt ohne methodische Artefakte zu erheben, spielen in Bezug auf die Items eines Tests auch Validitätsaspekte eine Rolle. In Kapitel 1.7 und 3.6 wurde auf der Grundlage des dort vorgestellten Entitätenrealismus Validität als Eigenschaft eines Tests aufgefasst, die angibt, ob ein Test auch diejenige Eigenschaft misst, die er zu messen vorgibt. Ein wesentliches Merkmal dieser Sichtweise war, dass die latente Variable kausal für die Antwort einer Versuchsperson auf einem Item verantwortlich ist. Dementsprechend muss ein wesentliches Ziel der Validierung eines Tests der Nachweis sein, dass die kausale Verbindung zwischen latenter Eigenschaft und Antwortverhalten existiert. Gefragt wird also nach dem kognitiven Prozess, welcher dem Antwortverhalten der Versuchspersonen zugrunde liegt. Die Frage nach der Validität eines Instruments ist somit keine Frage der Methoden, mit welcher ein Test validiert werden soll, sondern es handelt sich hierbei um eine substanzielle, psychologische Theorie darüber, wie das Antwortverhalten von Versuchspersonen zustande kommt. Gezeigt werden muss, wie Unterschiede in der zu messenden Eigenschaft jeweils ein unterschiedliches Antwortverhalten und damit verbunden die Wahl einer anderen Antwortkategorie verursachen (Borsboom et al., 2004, Embretson, 1983). Beispiele für ein solches Vorgehen zur Validierung gibt es hauptsächlich aus dem Bereich der

Leistungs- und Intelligenzdiagnostik sowie der Entwicklungsdiagnostik (Embretson, 1994, 1998, Jansen & van der Maas, 1997, 2002).

Um einen solchen Nachweis zu erbringen, sind mehrere Schritte notwendig. Zuerst muss aus der psychologischen Theorie ein formales Modell entwickelt werden, im hiesigen Kontext muss ein psychologisches Modell epistemologischer Überzeugungen in ein formales Modell epistemologischer Überzeugungen übertragen werden. Im zweiten Schritt muss aus diesem formalen Modell ein bestimmtes Antwortverhalten abgeleitet werden. Im dritten Schritt wird auf der Basis der Annahmen über das Antwortverhalten ein bestimmtes Antwortmuster abgeleitet, welches auch an empirischen Daten überprüft werden kann. Eine solche Überprüfung kann z. B. auf der Basis einer konfirmatorischen Faktorenanalyse oder auch der Item-Response-Theorie erfolgen, wobei die Messmodelle dem geforderten formalen Modell entsprechen, aus welchen sich die erwarteten Antwortmuster ableiten lassen. Der alleinige Einsatz statistischer Methoden ist aber nicht ausreichend, um Validität in dem hier geforderten Sinn nachzuweisen. So stellt z. B. die faktorielle Validität nur eine notwendige, aber keine hinreichende Bedingung für das Vorliegen von Konstruktvalidität im Sinne einer kausalen Beeinflussung der manifesten durch die latente Variable dar.

Die explizite Verbindung eines psychologischen mit einem formalen Modell sowie ein darauf beruhendes Modell des Antwortverhaltens ist in den bisher vorliegenden Studien zu epistemologischen Überzeugungen nicht vorhanden.⁵⁵ Es fehlt eine psychologische Theorie epistemologischer Überzeugungen, die angibt, in welcher Art und Weise unterschiedliche epistemologische Überzeugungen bzw. unterschiedliche Ausprägungen einer epistemologischen Überzeugung zu einem bestimmten Antwortverhalten bei den Items führen. Hier stellt sich das grundsätzliche Problem der kognitiven Psychologie, dass kein direktes Maß eines kognitiven Prozesses erhoben werden kann (Limón, 2006b). Komplizierter wird dies noch dadurch, dass nicht mit Sicherheit bestimmt werden kann, welche Grundlage eine Versuchsperson bei der Beantwortung eines Items zur Auswahl der Antwortkategorie heranzieht (Karabenik et al., 2007, Klopp & Stark, im Druck, Leach et al., 2000). Somit kann es möglich sein, dass beim Beantworten eines Items von zwei Versuchspersonen die gleiche Antwortkategorie gewählt

⁵⁵ Dies ist aber nicht nur ein Problem in der Forschung zu epistemologischen Überzeugungen, sondern allgemein ist in der Psychologie wenig erforscht, wie eine Eigenschaft mit einem bestimmten Antwortverhalten zusammenhängt (Sijtsma, 2012). Hauptsächlich im Bereich der Persönlichkeitspsychologie mangelt es an einer Theorie, wie beispielsweise eine Eigenschaft wie Extraversion mit dem Antwortverhalten auf Extraversion-Items zusammenhängt. Beispiele, wie der Zusammenhang von Eigenschaften und Antwortverhalten spezifiziert werden können stellt Kyngdon (2013) vor.

wird, obwohl bei der Itembeantwortung völlig unterschiedliche Antwortprozesse zugrunde liegen. Ebenso kann es möglich sein, dass zwei Personen mit einem gleichen Verständnis eines Items zwei verschiedene Antwortkategorien wählen.

Eine Möglichkeit, die sich hier ergibt, ist die Untersuchung der kognitiven Prozesse, die bei der Itembeantwortung ablaufen. Entsprechende Methoden finden bei der Konstruktion von Items in der Umfrageforschung Verwendung. Zusammengefasst handelt es sich hier um kognitive Pretest-Verfahren, mit denen geprüft wird, ob die Befragten die dem jeweiligen Item zugrunde liegende Intention richtig verstehen (vgl. Willis, 2005). Hierunter fällt eine Reihe spezieller Techniken, wie z. B. Nachfragen zum Verständnis eines Items oder Nachfragen zur Wahl einer Antwortkategorie. Eine andere Technik besteht in der Paraphrasierung eines Items, bei dem die Versuchsperson ein Item in ihren eigenen Worten wiedergeben muss, oder der Prozess der Itembeantwortung wird mithilfe der Methode des lauten Denkens erhoben. Diese Methoden gestatten einen Einblick, welche kognitiven Prozesse bei der Itembeantwortung bei den Versuchspersonen ablaufen können. Allerdings sind diese Methoden in erster Linie explorativ und nicht Teil einer psychologisch fundierten Theorie über den eigentlichen Antwortprozess. Allerdings bieten die Ergebnisse solcher Analysen die Möglichkeit, eine Theorie über den eigentlichen Antwortprozess zu entwickeln.

Eine Studie im Rahmen kognitiver Pretest-Verfahren wurde von Caramazza (2010) durchgeführt. In dieser Untersuchung wurden Items des EPI-Fragebogens im Hinblick auf das Antwortverhalten untersucht. In dieser Untersuchung wurden die Methode des lauten Denkens sowie eine Kombination des Paraphrasierens und der Nachfragetechnik angewandt. Beim lauten Denken wurden die ausgewählten Items des EPI-Fragebogens vorgelesen und die Versuchspersonen mussten laut aussprechen, durch welche Gedankengänge eine bestimmte Antwortalternative auf der vorgegebenen Ratingskala ausgewählt wurde. Vor der eigentlichen Erhebung wurde mit den Versuchspersonen ein Probedurchgang ausgeführt, um die Methode des lauten Denkens zu üben. Beim Paraphrasieren wurde der Versuchsperson zuerst das Item vorgelesen und anschließend wurde die Antwortkategorie von der Versuchsperson ausgewählt. Anschließend sollte die Versuchsperson das Item in eigenen Worten wiedergeben. Im Anschluss hieran wurden der Versuchsperson Verständnisfragen zu einzelnen Ausdrücken bzw. Wörtern aus dem Item gestellt, die zuvor aufgrund theoretischer Überlegungen als problematisch bestimmt worden waren. Auf diese Weise wurden jeweils zehn Versuchspersonen mithilfe des lauten Denkens und zehn Versuchspersonen mithilfe des Paraphrasierens in

Kombination mit der Nachfragetechnik untersucht. Insgesamt nahmen somit 20 Versuchspersonen (16 Frauen und 4 Männer) im Alter zwischen 23 und 32 an der Untersuchung teil, wobei 18 Probanden angaben, noch keine Erfahrung mit den beiden durchgeführten Methoden zu besitzen. Alle Versuchspersonen waren Studenten der Universität des Saarlandes, die entweder einem Lehramtsstudiengang oder einem Magisterstudiengang angehörten. Vorgelegt wurden 20 Items des EPI-Fragebogens, die aufgrund der beschriebenen Untersuchungen als problematisch eingestuft wurden. Mögliche erwartete Probleme ergeben sich z. B. durch die unterschiedliche Interpretation in den Items vorkommenden Begriffen oder durch einen von Versuchsperson zu Versuchsperson unterschiedlichen Kontext, auf welchen die Antwort auf ein Item bezogen wird. Es wurden den Probanden Items aus den Skalen *Reflexive Natur des Wissens*, *Absolutes Wissen*, *Soziale Komponente des Wissens*, *Value of Knowledge*, *Kulturspezifische Wissenszugänge* sowie *Lernen lernen* vorgegeben. Die Antworten wurden zuerst aufgezeichnet und anschließend transkribiert. Die auf diese Weise gewonnenen Daten wurden mithilfe der qualitativen Inhaltsanalyse (Mayring, 1996) ausgewertet. Durch diese qualitative Analyse des erhobenen Materials sollten erfasste Prozesse der Itembeantwortung aufgrund ihrer inhaltlichen Ähnlichkeit kategorisiert werden. Die gefundenen Kategorien beinhalten somit ähnliche Antwortprozesse und werden im Folgenden zur Differenzierung der Antwortkategorien der Ratingskala als Prozesskategorien bezeichnet. Die gebildeten Prozesskategorien wurden anschließend mit den gewählten Antwortkategorien kreuztabelliert. Untersucht werden sollte, ob die Wahl einer bestimmten Antwortkategorie der Ratingskala systematisch mit einer bestimmten Prozesskategorie zusammenhängt.

Die Auswertung ergab, dass sich die Antworten der Probanden je nach angewandter Methode der Datenerhebung in unterschiedliche Prozesskategorien einordnen lassen. Dieser Befund ist methodologisch interessant, muss aber für die hier interessierende Fragestellung nur in der Hinsicht beachtet werden, dass die Ergebnisse der Prozesskategorienbildung einer Methode nicht notwendigerweise auf die Prozesskategorienbildung der anderen Methode übertragen werden können bzw. vergleichbar sind. Die Ergebnisse von Caramazza (2010) lassen sich grob in zwei Muster einteilen. Das erste Muster entspricht der recht homogenen Wahl einer Antwortkategorie, die allerdings von unterschiedlichen Prozesskategorien abhängt. Das zweite Befundmuster ist eine heterogene Wahl der Antwortkategorien, die von unterschiedlichen Prozesskategorien abhängen. Das zweite Befundmuster ist für eine weiterführende Analyse ungeeignet, da sich in der Kreuztabelle nur zufällig zusammen auftretende Antwort- und Pro-

zesskategorien befinden und keine Aussagen über Zusammenhänge von Prozess- und Antwortkategorien getroffen werden können.

Das erste Befundmuster ist hingegen aufschlussreicher. Es zeigt, dass die Wahl einer Antwortkategorie bei verschiedenen Versuchspersonen von unterschiedlichen kognitiven Prozessen hervorgerufen wird. So wählten die Versuchspersonen bei dem Item „Die Beurteilung von Wissen verändert sich mit neuen Erfahrungen“ fast überwiegend die Antwortkategorie „Stimmt weitgehend“, die Wahl wurde aber von verschiedenen Prozesskategorien hervorgerufen. So fanden sich beim lauten Denken die Prozesskategorien *Bedeutung der Beurteilung*, *Persönlich-Wissenschaftlich*, *Beeinflussung durch Lernen und neue Erkenntnisse*, *Sozialisati-on – Entwicklung* sowie *Zwischenmenschliche Beziehungen*. Mit der Methode des Paraphrasierens ergaben sich die Prozesskategorien *Persönliche Erfahrung*, *Wertung von Wissen*, *Konkretes Beispiel*, *Austausch* und *Wissenschaftliche Erfahrung*. Bei dem Item „In der Wissenschaft werden verschiedene Phänomene der Welt objektiv erklärt“ ergab sich ebenfalls eine homogene Wahl der Antwortkategorien durch die Versuchspersonen. Lautes Denken führte hierbei zu den Prozesskategorien *Anspruch Wissenschaft*, *Methodik-Theorie*, *Intuition*, *Unbewusste Beeinflussung* und *Gegensatz Naturwissenschaften – Geisteswissenschaften*, das Paraphrasieren führte zu den Prozesskategorien *Allgemeinheit*, *Wissenschaftliche Überprüfung*, *Unvoreingenommenheit*, *Realitätsnähe*, *Neutral* sowie *Von vielen akzeptiert*.

Ein weiteres interessantes Ergebnis, das hier noch kurz erwähnt werden soll, ist der Kontextbezug, den die Versuchspersonen beim Antworten auf die vorgelegten Items zeigten. Die Lehramtsstudenten bezogen sich beim Beantworten der Items weitgehend auf Beispiele aus dem Alltag, während die Magisterstudenten sich mehr auf wissenschaftliche Literatur und Forschungsergebnisse bezogen. Dieses Ergebnis war überdies unabhängig von der verwendeten Methode des kognitiven Interviews.

Die Ergebnisse verdeutlichen, dass selbst bei der Wahl gleicher Antwortkategorien unterschiedliche kognitive Prozesse zu dieser Auswahl geführt haben. Die Probanden greifen somit auf unterschiedliche Interpretationen eines Items bzw. der darin vorkommenden Begriffe zurück, um zur Wahl einer Antwortkategorie zu gelangen. Dies ist konform mit der zuvor genannten Annahme von Leach et al. (2000), dass es nicht möglich ist zu kontrollieren, welche Grundlage bei der Beantwortung eines Items herangezogen wird. Allerdings zeigt die Prozesskategorienbildung, die mit den Antworten der Versuchspersonen vorgenommen werden

konnte, aber auch, dass die bei der Itembeantwortung ablaufenden Prozesse eine Reihe von Gemeinsamkeiten aufweisen. Die Ergebnisse legen aber nahe, dass die Items in den meisten Fällen wohl nicht das Konstrukt erfassen, das vom Konstrukteur des Items beabsichtigt wurde. Es ist insofern fraglich, ob mithilfe dieser Items die Erfassung epistemologischer Überzeugungen gelingt.

Für die dargestellte Sichtweise der Validität als Kausalzusammenhang des interessierenden Konstrukts mit der Beantwortung des Items (Borsboom et al., 2004, vgl. Kapitel 3.6.2) haben diese Befunde Relevanz. Die Ergebnisse der Studie von Caramazza (2010) deuten darauf hin, dass die Versuchspersonen bei der Beantwortung der Items eher idiosynkratische Interpretationen der Items vornehmen, als dass diejenige Eigenschaft erhoben wird, die von dem Entwickler der Items intendiert wurde. Zwar kann durch solche explorative Studien nicht der geforderte Nachweis erbracht werden, dass die zu messende Eigenschaft kausal das Zustandekommen des Antwortverhaltens bedingt, da bei dieser Art von Untersuchung die zur Kausalanalyse notwendigen Manipulation nicht vorhanden ist (Steyer, 2004). Solche explorative Studien bieten jedoch erste Einblicke in die der Itembeantwortung zugrunde liegenden kognitiven Prozesse. Sie sind somit ein erster Schritt in die von Borsboom et al. (2004) geforderte substanzielle psychologische Theorie des Antwortverhaltens zu einem Messinstrument. Hierzu ist es aber zum einen erforderlich, eine psychologische Theorie epistemologischer Überzeugungen vorzulegen, die es zu bestimmen gestattet, in welcher Art und Weise die Items eines Fragebogens kognitiv verarbeitet werden. Zum anderen bedarf es einer Theorie, die angibt, in welcher Weise ein Item zu einem bestimmten Antwortverhalten führt. Neben diesen substanziell-psychologischen Arbeiten sind zur Validierung die entsprechenden methodischen Arbeiten erforderlich.

Bromme et al. (2008) stellen eine Theorie vor, welche die Interpretation eines Items zur Erfassung einer epistemologischen Überzeugung als Prozess des Textverstehens mithilfe des *Construct-integration*-Modells von Kintsch (1988) ansieht, die in Kapitel 6.2.3 genauer vorgestellt wird. Neben einer solchen Theorie ist es außerdem erforderlich, entsprechende Annahmen zu formulieren, wie eine Person zur Wahl der Antwortkategorie für dieses Item gelangt. Diese Theorien stellen damit die psychologischen Theorien bei der Entwicklung des Messinstrumentes dar und müssen mithilfe entsprechender Methoden überprüft werden. Ergänzend hierzu können die statistischen Eigenschaften mithilfe der Methoden, wie sie z. B. im Rahmen der vorliegenden Arbeit durchgeführt worden, geprüft werden. Die statistischen Ei-

genschaften beziehen sich somit auf das formale Modell. Die Prüfung der Validität eines Messinstruments kann sich folglich nicht nur auf die Prüfung von statistischen Eigenschaften und Passung der Messmodelle beschränken.

Eine Möglichkeit zur Bestimmung der Validität, die sich nicht auf eine Prüfung der Passung von Messmodellen beschränkt, hat Bornstein (2011) mit dem prozessorientierten Validitätsmodell vorgeschlagen. Im prozessorientierten Validitätsmodell wird Validität konzeptualisiert als das Ausmaß, in dem sich ein spezifizierter Prozess zwischen Itemdarbietung und -beantwortung nachweisen lässt. Bei der Messung epistemologischer Überzeugungen wäre daher ein Prozess zu spezifizieren, bei dem die Ausprägung einer bestimmten epistemologischen Überzeugung zur Wahl einer bestimmten Antwortkategorie führt. Ist dieser Prozess nachgewiesen, wird versucht, diesen mittels gezielter, experimenteller Variation zu verändern. Eine Möglichkeit zur Anwendung des prozessorientierten Validitätsmodells bei der Validierung von Skalen zur Erfassung epistemologischer Überzeugungen haben Klopp und Stark (im Druck) aufgezeigt. Zur Prüfung der Validität wird hierbei ein zweistufiges, an das Vorgehen von Karabenick et al. (2007) angelehntes Verfahren verwendet. Der Ansatzpunkt des Verfahrens ist wie bei Caramazza (2010) die Methode der kognitiven Interviews, wobei sich das Verfahren in zwei Schritte aufteilt. In einem ersten Schritt muss untersucht werden, inwieweit bei verschiedenen Probanden jeweils die gleiche Prozesskategorie bei der Itembeantwortung herangezogen wird. Die Prozesskategorie muss dem zu messenden Konstrukt entsprechen, so dass Unterschiede in der Itembeantwortung lediglich durch Unterschiede in dem zugrunde liegenden Konstrukt bedingt werden und nicht durch unterschiedliche Konstrukte. Zeigen sich unterschiedliche Prozesskategorien und liegen der Itembeantwortung daher unterschiedliche Konstrukte zugrunde, werden interindividuelle Unterschiede und durch verschiedene Konstrukte bedingte Unterschiede bei der Itembeantwortung konfundiert. Als Ergebnis des ersten Schritts werden die Items bestimmt, die gleiche Prozesskategorien aufweisen. Für diese Items muss im zweiten Schritt untersucht werden, inwieweit bei denselben zugrunde liegenden Konstrukten die Antwortkategorien der Ratingskala kohärent verwendet werden. Erfolgt die Verwendung der Antwortkategorien inkohärent, werden interindividuelle Unterschiede und durch inkohärente Verwendung der Antwortkategorien bedingte Unterschiede konfundiert. Für diese Analysen müssen die Items herangezogen werden, die am Ende des ersten Schritts ausgewählt wurden. Für jeden Probanden muss dann geprüft werden, ob sich bei gleicher Prozesskategorie auch eine kohärente Wahl der Antwortkategorien erkennen lässt. Im Anschluss an die Itemauswahl kann mittels Interventionsstudien geprüft werden, ob sich der

Antwortprozess durch eine experimentelle Manipulation verändern lässt (Mislevy, 2008). Beispielsweise kann mittels eines Priming-Experimentes untersucht werden, ob sich der Antwortprozess im Sinne der Primingmaßnahme verändern lässt. Die experimentelle Manipulierbarkeit würde somit einen Nachweis für den von Borsboom, Mellenbergh und van Heerden (2004) geforderten Kausalzusammenhang zwischen Konstrukt und Itembeantwortung darstellen. Ein hypothesenkonformer Befund entspräche dann dem spezifizierten Prozess und würde für die Validität des Verfahrens sprechen.

Diese Anmerkungen können nur als Anregung dienen, wie Prüfungen der Validität von Instrumenten, die auf einer psychologischen Theorie basieren, aussehen können. Insbesondere werden ausgearbeitete und prüfbare Theorien der Item-Antwortprozesse benötigt. Aber auch die eingangs erwähnten Aspekte der Itemgestaltung müssen berücksichtigt werden, um zu anwendbaren Messmodellen zu gelangen. Neben den diskutierten Aspekten der Itemgestaltung sowie der Validität gibt es aber noch die Problematik der Reaktivität der Erhebung epistemologischer Überzeugungen, die im nächsten Kapitel besprochen werden soll.

6.2.3 Reaktivität der Itembeantwortung

Es hat sich gezeigt, dass oftmals die im Verhalten sichtbar werdenden epistemologischen Überzeugungen von denen abweichen, die sich mittels Fragebogen oder Interviewverfahren erheben lassen (vgl. Bell & Linn, 2002, Leach et al., 2000, Olafson & Schraw, 2006). Hierbei unterscheidet man zwischen erklärten und gezeigten epistemologischen Überzeugungen (Louca et al., 2004, Limón, 2006b). Erklärte epistemologische Überzeugungen entsprechen den epistemologischen Überzeugungen, wie sie aus Fragebogen oder Interviews abgeleitet und einer Person zugeschrieben werden. Gezeigte epistemologische Überzeugungen entsprechen den epistemologischen Überzeugungen, die ein Beobachter aus dem im Lernen oder Unterrichten gezeigten Verhalten ableiten und einer Person zuschreiben würde.

Olafson und Schraw (2006) untersuchten den Zusammenhang zwischen epistemologischen Überzeugungen und den Unterrichtspraktiken von Lehrkräften. Dazu wurde von Lehramtsstudierenden als Versuchspersonen das EBI von Schraw et al. (2002) bearbeitet. Hieraus wurde abgeleitet, ob die Versuchsperson als Realist, Kontextualist oder Relativist einzustufen ist. Realisten glauben, dass es eindeutiges und stabiles Wissen gibt, das am besten durch Experten vermittelt wird. Kontextualisten gehen davon aus, dass Wissen immer in den Kontext des Lernens eingebettet und veränderbar ist, die Wissensvermittlung erfolgt in der gemeinsamen

Konstruktion von Wissen, bei welchem der Lehrer als Vermittler dient. Relativisten nehmen an, dass Wissen subjektiv und leicht veränderbar ist und dass jeder Lerner sein eigenes Verständnis bildet, das nicht mit dem Verständnis anderer Lerner übereinstimmen muss. Zusätzlich zu dem EBI wurden die Lehramtsstudierenden zu ihren Unterrichtspraktiken interviewt.

Die Mehrheit der Versuchspersonen war als Kontextualisten zu bezeichnen, gefolgt von dem relativistischen Typ, wobei sich allerdings keine klare Trennung zwischen den drei Typen aufrechterhalten ließ. Die Mehrheit der Versuchspersonen wiesen auch mehr als einen Typus auf. Es erwies sich weiterhin, dass die gefundenen Typen epistemologischer Überzeugungen inkongruent zu dem gezeigten Lehrverhalten waren. Obwohl kontextualistische bzw. relativistische Sichtweisen vorherrschten, wurden im Lehralltag Verhaltensweisen gezeigt, die einer realistischen Sichtweise entsprechen. Dies drückt sich in der Verwendung von Unterrichtsmethoden aus, die auf eine reine Vermittlung der Inhalte ausgelegt sind.

Die mittels Fragebogen erfassten epistemologischen Überzeugungen stimmen demzufolge nicht mit dem Verhalten überein, das sich Personen in Selbstbeurteilungen zuschreiben. Bei der Erfassung epistemologischer Überzeugungen durch Befragung stellt sich das Problem der Reaktivität von Fragebogenerhebungen. Reaktivität bedeutet, dass das gewählte Messverfahren das zu messende Phänomen verändert. Reaktivität tritt vor allem dann auf, wenn das zu messende Phänomen den Befragten fremd und nicht Teil ihrer normalen Umwelt ist (Campbell, 1957, Petermann & Noack, 1984). Im Normalfall ist es einer Person nicht bewusst, dass sie epistemologische Überzeugungen besitzt (Limón, 2006b). Wird nun die Beurteilung eines Items verlangt, löst dies einen kognitiven Prozess aus, in welchem der Inhalt des Items verarbeitet und eine Antwortalternative bestimmt wird (Tourangeau et al., 2000). Dieser kognitive Prozess kann als „Unschärferelation“ angesehen werden, epistemologische Überzeugungen kommen nicht direkt zum Vorschein, sondern kommen nur reflektiert durch die Person zum Vorschein, da der ausgelöste kognitive Prozess wie ein Filter wirkt (vgl. Gubler & Bischof, 1993).

Das Problem der Reaktivität lässt sich in das Prozessmodell der Itembeantwortung von Podsakoff, MacKenzie, Lee und Podsakoff (2003) einbetten. Nach diesem Modell lässt sich der Prozess der Itembeantwortung in die aufeinanderfolgenden Stadien Verständnis, Abruf, Urteil, Antwortwahl und Antwortabgabe einteilen. Reaktivität ist bei den Stadien Abruf und Urteil von Relevanz, weil die zur Itembeantwortung notwendigen Informationen aus dem

Gedächtnis abgerufen werden müssen. Da epistemologische Überzeugungen in der Regel weit von den Alltagserfahrungen des Antwortenden entfernt sind, müssen beim Abruf der Informationen aus dem Gedächtnis entsprechende Anstrengungen unternommen werden, sofern die notwendigen Informationen überhaupt vorhanden sind und nicht erst konstruiert werden müssen.

Dieser Prozess des Informationsabrufs selbst ist nicht in dem Modell von Podsakoff et al. (2003) enthalten. Da es sich bei den Items zur Erfassung epistemologischer Überzeugungen um Aussagesätze handelt, haben Bromme et al. (2008) vorgeschlagen, den Informationsabruf beim Beantworten eines Fragebogenitems mithilfe des *Construct-integration*-Modells (CI-Modell) von Kintsch (1988, 1998) zu beschreiben. Nach dem CI-Modell wird das zum Beantworten eines Items notwendige Wissen in Form eines präpositionalen Netzwerks abgespeichert und beim Lesen des Items wird dieses Wissen aktiviert, wobei sich zwei Phasen unterscheiden lassen: die *Konstruktionsphase* und die *Integrationsphase*. Während der Konstruktionsphase werden beim Lesen die präpositional repräsentierten Inhalte, die zum Verständnis des Items notwendig sind, aktiviert. Darüber hinaus können während der Konstruktionsphase Schlussfolgerungen, die sich aus dem Iteminhalt ergeben, sowie Generalisierungen des Iteminhalts gebildet werden. Das Ergebnis der Konstruktionsphase stellt ein kohärentes Netzwerk aller aktivierten Inhalte dar. In der anschließenden Integrationsphase werden diese aktivierten Inhalte in ein Situationsmodell des Textes überführt. Dabei werden zu dem kohärenten Netzwerk zusätzliche Assoziationen aus dem Gedächtnis abgerufen. Diese Gedächtnisabrufe sind abhängig vom jeweiligen Wissen eines Probanden, der ein Item bearbeitet. Das CI-Modell geht davon aus, dass der einen Satz umgebende Text Randbedingungen für das Verständnis schafft. Bei dem Item eines Fragebogens liegen diese Randbedingungen allerdings nicht in Form eines Textes vor. Vielmehr wird davon ausgegangen, dass das Umfeld sowie das bisherige Wissen einer Person die Randbedingungen für das Verständnis eines Items darstellen. Als Abschluss der Konstruktionsphase liegt ein Situationsmodell des Items vor. Im Situationsmodell sind der Inhalt des Items sowie die während der Konstruktionsphase entstandenen Assoziationen enthalten.

Im Fall der Items zur Erfassung epistemologischer Überzeugungen kann davon ausgegangen werden, dass die abgefragten Inhalte nicht zum Alltagskontext einer Person gehören und ihr, wie bereits angeführt wurde, nicht bewusst ist, dass sie über epistemologische Überzeugungen

verfügt.⁵⁶ Daher müssen in der Regel während der Konstruktionsphase entsprechende Assoziationen hergestellt werden, um zu einem Situationsmodell des Items zu gelangen. Dadurch ist es aber auch möglich, dass der Proband in der Konstruktionsphase nach Inhalten suchen muss, die nicht direkt im Zusammenhang mit epistemologischen Überzeugungen stehen.

In der Abrufphase nach Podsakoff et al. (2003) kann es also vorkommen, dass Inhalte aus dem Gedächtnis abgerufen werden, die nicht im Zusammenhang mit epistemologischen Überzeugungen stehen. Ebenso kann die Relevanz der abgerufenen Inhalte während der Urteilsphase falsch eingeschätzt werden. Beide Faktoren können dazu führen, dass bei den Selbsteinschätzungen nicht mit dem tatsächlichen Verhalten übereinstimmende Antworten abgegeben werden. Dass Selbsteinschätzungen immer einer gewissen Subjektivität durch den Antwortenden unterliegen, ist in der Fragebogenforschung schon länger thematisiert worden (Mummendey, 1999). Die Antworten auf Items schließen demnach auch immer die subjektive Interpretation des Items ein. Auch wenn Items sich auf konkretes Verhalten beziehen, was den einfachsten Fall der Selbstbeurteilung darstellt, muss das Item verstanden und von den Probanden verarbeitet werden (Fisseni, 2004). Schwieriger gestaltet sich die Itembeantwortung bei einem für die Versuchsperson unbekanntem Sachverhalt wie epistemologische Überzeugungen.

In den ersten Kapiteln der Diskussion wurden methodische bzw. methodologische Aspekte der Erfassung epistemologischer Überzeugungen mittels Fragebogen diskutiert. Nach der Abgrenzung epistemologischer Überzeugungen von anderen Konstrukten standen Validitätsfragen und die Reaktivität bei der Erfassung epistemologischer Überzeugungen im Mittelpunkt. Im nächsten Kapitel werden weitere Probleme bei der Erfassung epistemologischer Überzeugungen diskutiert, wobei aber inhaltliche Fragen in den Mittelpunkt rücken. Als erster inhaltlicher Aspekt wird im nächsten Kapitel die normative Sichtweise epistemologischer Überzeugungen besprochen.

⁵⁶ Hier sollte der Begriff des Alltagskontexts klargestellt werden, vor allem im Hinblick auf die im letzten Kapitel vorgestellten Ergebnisse von Caramazza (2010), wonach Lehramtsstudierende bei der Beantwortung von Items auf Alltagszenarien zurückgreifen. Im Zusammenhang mit dem Problem der Reaktivität bedeutet Alltagsferne, dass die Beurteilung epistemologischer Überzeugungen üblicherweise nicht von sich aus durchgeführt wird, sondern erst dann ins Bewusstsein der Probanden rückt, wenn diese, wie z. B. durch einen Fragebogen, darauf aufmerksam gemacht werden. Dass bei der eigentlichen Beurteilung eines Items dann Beispiele aus dem Alltag durch einen Probanden aufgegriffen werden, ist ein anderer Sachverhalt.

6.3 Normative Sichtweise epistemologischer Überzeugungen

6.3.1 Produktivität, Korrektheit und Kontextabhängigkeit

In der Literatur wird häufig eine normative Sichtweise epistemologischer Überzeugungen vertreten. Am Beispiel der Dimension *Sicherheit des Wissens* verdeutlicht bedeutet die normative Sichtweise, dass die Überzeugung einer Person eine fortgeschrittene epistemologische Überzeugung ist, wenn diese denkt, dass Wissen mehr tentativ als sicher ist. Durch ein entsprechendes gepoltes Item in einem Fragebogen abgefragt, hätte diese Person einen hohen Wert auf der Ratingskala erreicht. Personen, die Wissen als sicher ansehen, hätten einen niedrigen Wert auf der Skala und ihre epistemologische Überzeugung würde als naiv bezeichnet werden.

Diese Sichtweise wird von Elby und Hammer (2001) infrage gestellt. Nach Elby und Hammer ist weniger die Ausprägung einer epistemologischen Überzeugung entscheidend, um diese als fortgeschritten zu bezeichnen, als vielmehr die Produktivität und Korrektheit. Die Produktivität bedeutet, dass eine epistemologische Überzeugung zu einem Ergebnis führt, das im Vergleich mit einem objektiven Kriterium übereinstimmt oder zu einem Ergebnis führt, das den Erfordernissen einer Situation entspricht. Die Produktivität einer epistemologischen Überzeugung setzt somit ein Vergleichskriterium voraus. Korrektheit einer epistemologischen Überzeugung bedeutet, dass die epistemologische Überzeugung einer Person den Vorgaben einer Gemeinschaft wie z. B. einer Scientific Community entspricht. Die Begriffe Produktivität und Korrektheit werden unten an Beispielen verdeutlicht. An dieser Stelle wird allerdings erkennbar, dass Produktivität und Korrektheit auch vom jeweiligen Kontext abhängen. Der Kontext kann einerseits in dem Sachverhalt, der dem Vergleichskriterium zugrunde liegt, andererseits in den Vorgaben der jeweiligen Gemeinschaft bestehen.

Nach Elby und Hammer (2001) sind die meisten der Fragebogenitems nicht dazu geeignet, die Produktivität und Korrektheit epistemologischer Überzeugungen zu erfassen. Die meisten Items stellen demnach pauschale Verallgemeinerungen dar, welche den Kontext, in dem eine epistemologische Überzeugung eine Rolle spielt, verallgemeinert. Diese Sichtweise wird im Folgenden an einigen Beispielen verdeutlicht, die drei der vier Dimensionen von Hofer und Pintrich (1997) entsprechen, nämlich *Sicherheit des Wissens*, *Wissensquelle* und *Komplexität*

des Wissens.⁵⁷ Gleichzeitig werden die Begriffe Produktivität und Korrektheit beispielhaft erläutert.

Ein Item zur Erfassung der Dimension *Sicherheit des Wissens* ist z. B. „Nichts außer dem Tod ist sicher.“ Die normative Sichtweise schreibt einer Person eine fortgeschrittene epistemologische Überzeugung zu, die Wissen als vorläufig und nicht sicher ansieht. Eine solche Person würde der Aussage des Items also in hohem Maße zustimmen. Diese Zustimmung ist aber weder produktiv noch korrekt, was nun am fiktiven Beispiel eines Lerner, der die Newton'sche Gesetze der Bewegung in einem Einführungskurs Physik lernt, illustriert wird.

Für diesen Lerner ist es nicht produktiv, diese Bewegungsgesetze als nicht sicher anzusehen. Da aber insbesondere das 2. Newton'sche Gesetz teilweise dem allgemeinen Verständnis widerspricht, kann es möglich sein, dass Personen, welche die Bewegungsgesetze lernen, genau dieses Gesetz nur für die gegebenen Beispiele als sicher ansehen und anwenden und für andere Kontexte nicht. Bei gestellten Aufgaben wird dann auf das Alltagsverständnis ausgewichen, da das 2. Bewegungsgesetz intuitiv eher schwer verständlich ist und dem Alltagsverständnis widerspricht. Sieht ein Lerner nun dieses Gesetz als tentativ an und wendet sein intuitives Alltagswissen an, kann es vorkommen, dass falsche Schlussfolgerungen gezogen bzw. gestellte Aufgaben falsch gelöst werden. Dies konnte von diSessa (1993, zitiert nach Elby & Hammer, 2001) nachgewiesen werden. Im Kontext einer Klausur kann dies dazu führen, dass eine Aufgabe falsch gelöst wird. Die Überzeugung, dass die Bewegungsgesetze nicht sicher sind, ist somit nicht produktiv, weil die Übereinstimmung des Ergebnisses mit der Lösung der Aufgabe nach dem 2. Bewegungsgesetz als objektivem Vergleichskriterium fehlt.

Eine generelle Überzeugung, dass Wissen nicht sicher sein kann, ist nicht nur unproduktiv, sondern auch nicht korrekt. Z. B. spricht es nicht für eine fortgeschrittene epistemologische Überzeugung, wenn man es als unsicher ansieht, dass die Erde die Form einer Kugel hat. Allerdings ist die Festsetzung, ob eine epistemologische Überzeugung als naiv oder fortgeschritten anzusehen ist, auch abhängig vom Kontext. So ist das eben genannte Beispiel mit der Kugelgestalt der Erde im Kontext des Schulunterrichts oder einer universitären Einführungsveranstaltung wenig fortgeschritten. Im Kontext eines Seminars über Erkenntnistheorie, in dem die Möglichkeit, überhaupt zu einer sicheren Erkenntnis zu gelangen, diskutiert wird, kann

⁵⁷ Obwohl die Dimension *Komplexität des Wissens*, wie oben aufgeführt, keine epistemologische Überzeugung darstellt, wird dieses Beispiel von Elby und Hammer (2001) dennoch hier besprochen, um den grundlegenden Gedanken zu verdeutlichen.

das Zweifeln an der Kugelgestalt der Erde aber als fortgeschrittene epistemologische Überzeugung angesehen werden. Im Fall des Schulunterrichts oder der universitären Einführungsveranstaltung wäre die epistemologische Überzeugung an der Kugelgestalt der Erde zu zweifeln, nicht korrekt, weil in diesem Kontext üblicherweise diese Erkenntnis als gegeben und nicht änderbar angesehen wird. Im Kontext des Seminars über Erkenntnistheorie wird üblicherweise aber erwartet, dass man an der Möglichkeit einer sicheren Erkenntnis überhaupt zweifelt.

Bei der Dimension *Wissensquelle* ist die normative Sichtweise, dass eine epistemologische Überzeugung dann als fortgeschritten angesehen wird, wenn eine Person Wissen als internal gerechtfertigt und konstruiert ansieht. Die Übernahme des Wissens von Autoritäten wird als naiv angesehen. In dem Fragebogen von Schommer wird diese Dimension z. B. durch das Item „Manchmal muss man die Antworten des Dozenten akzeptieren, auch wenn man sie nicht versteht“ erfasst. Eine Person mit einer fortgeschrittenen epistemologischen Überzeugung würde dieser Aussage nur wenig zustimmen.

Diese normative Sichtweise ist nicht notwendigerweise produktiv. Hammer (1989) gibt ein fiktives Fallbeispiel eines Studenten in einer Einführungsveranstaltung Physik. Der Student besaß die Überzeugung, dass Wissen internal gerechtfertigt wird, im normativen Sinn somit also eine fortgeschrittene epistemologische Überzeugung. In der Veranstaltung lag der Schwerpunkt aber auf der Vermittlung algorithmischen Wissens (also der rechnerischen Lösung von physikalischen Problemen), ohne dass auf physikalisches Verständnis Wert gelegt wurde. Durch das genaue Elaborieren der Vorlesungsinhalte statt des simplen Lernens der Rechenwege geriet der Student mit den Veranstaltungsinhalten immer mehr in Verzug und konnte ihr deshalb nicht mehr folgen und entsprechend gering waren die Leistungen des Studenten in der Abschlussklausur des Kurses. Die normativ fortgeschrittene epistemologische Überzeugung des Studenten war somit nicht den Erfordernissen des Kurses angepasst und somit nicht produktiv.

Bei der Dimension *Komplexität des Wissens* bedeutet die normative Sichtweise, dass eine Person eine fortgeschrittene epistemologische Überzeugung besitzt, wenn Wissen als Vernetzung von Aussagen angesehen wird. Im Fragebogen von Schommer wird dies z. B. durch das Item „Ein Satz hat wenig Bedeutung, solange man den Kontext, in dem er gesprochen oder geschrieben wurde, nicht kennt“, zu erfassen versucht. Auch diese Überzeugung muss nicht

produktiv sein, als Beispiel kann der oben genannte Student in der Einführungsveranstaltung zur Physik dienen. Der Kontext des Kurses hat ein Auswendiglernen der Formeln und Rechenwege verlangt, nicht das Verstehen der physikalischen Zusammenhänge. Die Lernweise des Studenten, Zusammenhänge zu bilden, ging an den Erfordernissen des Kurses vorbei, die aus einer Abfrage der Formeln und Rechenweg bestand. Die epistemologische Überzeugung des Studenten war somit im Hinblick auf die vom Dozenten geforderten Ansprüche nicht produktiv.

Diese Überzeugung, dass Wissen aus vernetzten Bestandteilen besteht, muss auch nicht korrekt sein. Nach Elby und Hammer (2001) ist es angebracht, in den meisten Kontexten Wissen eher als Netz von verbundenen Aussagen als eine Sammlung einzelner Fakten anzusehen. Diese Sichtweise ist aber abhängig vom Kontext. So kann es im Kontext naturwissenschaftlichen Unterrichts sinnvoll sein, die Zusammenhänge der einzelnen Aussagen, Gesetze und Theorien etc. als zusammenhängend anzusehen, beim Lernen der Namen von Hauptstädten oder Flüsse eines Landes im Rahmen des Geografieunterrichts ist dies wenig förderlich, da es sich hier einfach um Faktenwissen handelt. Für das Wissen über die Hauptstädte von Ländern ist die fortgeschrittene epistemologische Überzeugung somit nicht korrekt.

Elby und Hammer (2001) konnten mit diesen Beispielen darlegen, dass die normative Sichtweise epistemologischer Überzeugungen nicht haltbar ist, und dass es vom Kontext abhängt, ob eine epistemologische Überzeugung als fortgeschritten oder naiv bezeichnet werden kann. Diese Kontextabhängigkeit wird von den Items der üblichen Fragebogen allerdings nicht erfasst. Als Ergebnis der Analyse von Elby und Hammer kann festgehalten werden, dass die Kontextabhängigkeit ein wesentliches Element ist, das epistemologische Überzeugungen charakterisiert und das nicht mit den üblichen Items erfasst werden kann.

Als erste Konsequenz sprechen sich Elby und Hammer (2001) dafür aus, diese Kontextabhängigkeit mithilfe von Interviewverfahren zu erfassen, weil in diesen im Gegensatz zu Fragebogen die jeweiligen Besonderheiten des Kontexts erfasst werden können. Interviewverfahren bieten den Vorteil, dass mit ihnen das phänomenale Erleben des Interviewten erfasst werden kann. Das phänomenale Erleben beinhaltet nach Bischof (2008), dass das Erleben des Interviewten in einer deskriptiven Sprache geschildert wird. Die im Interview gesammelten Daten müssen anschließend in ein funktionales Modell übertragen werden, das von der reinen

Deskription der Versuchspersonen absieht und eine erklärende Interpretation der Zusammenhänge von epistemologischen Überzeugungen und den Kontextvariablen gibt.

Die Kontextabhängigkeit von Erleben und Verhalten ist nicht nur im Rahmen der Erforschung epistemologischer Überzeugungen von Interesse, sondern ist ein Merkmal, welches für viele Bereiche der Psychologie kennzeichnend ist. Der Kontext bedingt das Verhalten und Erleben von Personen. Hierauf hat auch Graumann (2000) hingewiesen. Nach Graumann ist das Erleben und Verhalten einer Person nicht nur in einen Kontext eingebettet, vielmehr steht es auch in einem funktionalen Zusammenhang mit dem jeweiligen Kontext. Daraus ergibt sich, dass epistemologische Überzeugungen mit dem jeweiligen Kontext bzw. der Situation in einem funktionalen Zusammenhang stehen.

Die Abhängigkeit von Kontextvariablen hat Hammer und Elby (2002) dazu geführt, die Unitaritätsannahme für epistemologische Überzeugungen infrage zu stellen (vgl. Kapitel 1.4), da diese nicht mit einer Kontextabhängigkeit vereinbar ist. Fortgeschrittene epistemologische Überzeugungen erlauben nach Bromme et al. (2008) die kontextsensitive Beurteilung und Rechtfertigung von Aussagen. Diese kontextsensitive Beurteilung stellt den funktionalen Bezug zwischen Kontext und epistemologischer Überzeugung her. Dabei spielen aber weitere Faktoren – wie das themenbezogene Wissen und das ontologische Wissen – eine Rolle, um in einem gegebenen Kontext zu einem epistemologischen Urteil über die Rechtfertigung einer Aussage zu gelangen. Die Fähigkeit, in einem bestimmten Kontext eine flexible Beurteilung einer Aussage vornehmen zu können, soll die Diagnose einer Person als epistemologisch fortgeschritten oder naiv bestimmen. Dabei muss aber von der Unitaritätsannahme abgesehen werden, da diese bedeuten würde, dass in jedem Kontext das gleiche epistemologische Urteil gefällt würde. Dies kann nach der obigen Argumentation nicht der Fall sein kann. Bei einem solchen Rechtfertigungsprozess kommt somit immer auch intraindividuelle und intrakontextuelle Varianz zustande.

Bromme et al. (2008) verdeutlichen ihre Sichtweise am Beispiel einer Studentin, die sich im Rahmen eines Biologiestudiums mit dem kardiovaskulären System befassen muss. Dabei besucht sie zwei Kurse. Im ersten Kurs werden das kardiovaskuläre System und der Einfluss der Ernährung besprochen und entsprechende Laborübungen durchgeführt. Der zweite Kurs beschäftigt sich mit Forschungsmethoden und der Geschichte des Fachs Biologie. Die Studentin bemerkt nun einige Widersprüche zwischen beiden Kursen. Im ersten Kurs werden das kardi-

ovaskuläre System, dessen Bestandteile und dessen Funktionsweise als gegeben und die Aussagen und die Funktionsweise als sicher angesehen. Diese Haltung wird überdies verstärkt durch die Laborübungen, in denen das kardiovaskuläre System durch Messungen, Experimente etc. erfahrbar gemacht wird. In dem Kurs über Forschungsmethoden und die Geschichte der Biologie hingegen wird dargestellt, wie sich die Vorstellung über das kardiovaskuläre System im Laufe der Zeit entwickelte und wie für sicher gehaltenes Wissen revidiert wurde.

In diesem Beispiel würde man der Studentin eine fortgeschrittene epistemologische Überzeugung zusprechen, wenn diese die beiden Kurse differenziert betrachten kann und die Rechtfertigung von Aussagen im jeweiligen Kontext durchführt. Diese kann aber auch im Hinblick auf die beschriebene Produktivität epistemologischer Überzeugungen im Sinn von Hammer und Elby (2001) analysiert werden. Die Erkenntnisse über das kardiovaskuläre System infrage zu stellen, würde die tägliche Arbeit im Labor eher behindern als fördern, wäre also nicht produktiv (vgl. Bromme et al., 2008). Außerdem kann es je nach Gestaltung des Kurses und der anschließenden Klausur erforderlich sein, bestimmte Fakten über das kardiovaskuläre System nur reproduzieren zu können, unabhängig davon, ob der Inhalt verstanden wurde.

6.3.2 Normativität und die Entwicklung epistemologischer Überzeugungen

Die vorgebrachten Einwände gegen eine normative Sichtweise greifen aber noch weiter. Es gibt empirische Belege, dass mit zunehmender Beschäftigung mit einem bestimmten Fach die epistemologischen Überzeugungen der Lernenden naiver werden. Damit wird gleichzeitig aber auch die Entwicklung zu immer weiter fortgeschrittenen epistemologischen Überzeugungen, wie diese z. B. in den Stufenmodellen des entwicklungspsychologischen Ansatzes zu finden sind und als normative Vorstellung über die Entwicklung epistemologischer Überzeugungen angesehen werden (vgl. Kapitel 1.2), infrage gestellt.

Kienhues, Bromme und Stahl (2008) konnten experimentell im Rahmen einer Studie zur Veränderung von epistemologischen Überzeugungen nachweisen, dass mit zunehmendem Wissen in einer Domäne die epistemologischen Überzeugungen naiver werden. Hierzu wurden die epistemologischen Überzeugungen von Personen im Rahmen eines Vortests erhoben und anschließend wurden diese Personen in zwei Gruppen eingeteilt, je nachdem ob die Person im Vortest naive oder fortgeschrittene epistemologische Überzeugungen aufwies. Die epistemologischen Überzeugungen wurden mithilfe des CAEB (Stahl & Bromme, 2007) sowie des DEBQ (Hofer, 2000) erhoben. Die Versuchspersonen wurden anschließend zwei experimen-

tellen Bedingungen zugewiesen. Eine Gruppe erhielt einen Lehrtext mit Argumenten und Gegenargumenten zum Thema genetischer Fingerabdruck, die zweite Gruppe erhielt einen rein informativen Lehrtext zu demselben Thema. Aus dem vorgelegten Fragebogen wurden mittels einer Faktorenanalyse jeweils zwei Dimensionen extrahiert. Für die mit dem DEBQ erhobenen epistemologischen Überzeugungen zeigte sich, dass alle Gruppen bis auf diejenige mit naiven epistemologischen Überzeugungen und dem Lehrtext mit Argumenten und Gegenargumenten, nach der Intervention naivere epistemologische Überzeugungen als vorher aufwiesen. Insbesondere konnte aber nachgewiesen werden, dass die Gruppe mit fortgeschrittenen Überzeugungen, die den einfachen Lehrtext erhalten hatte, nach der Intervention naive epistemologische Überzeugungen in den DEBQ-Dimensionen aufwies. Tabelle 120 bietet eine Übersicht über die tendenziellen Veränderungen der jeweiligen Gruppen.

Tabelle 120: Tabelle mit den tendenziellen Änderungen aus Kienhues et al. (2008)

	DEBQ		CAEB	
	Stabilität	Einfachheit und Bestimmtheit	Textur	Variabilität
Gruppe mit informativem Text				
Naive	keine	keine	fortgeschrittener	fortgeschrittener
Fortgeschritten	naiver	naiver	keine	keine
Gruppe mit Gegenargumenten				
Naive	naiver	naiver	keine	keine
Fortgeschritten	naiver	naiver	naiver	naiver

Ähnliche Befunde konnten Köller, Baumert und Neubrand (2000) präsentieren, die einen Zusammenhang zwischen der Kurswahl in gymnasialen Oberstufen und epistemologischen Überzeugungen fanden. Dazu wurden die epistemologischen Überzeugungen mithilfe des EBQ erhoben. Die Schüler, die einen Leistungskurs in Physik wählten, wiesen nach Abschluss der Kurse eine ausgeprägtere dualistische Sichtweise als diejenigen Schüler auf, die nur einen Grundkurs gewählt hatten. Überdies stimmten die Schüler aus den Leistungskursen auch mehr der Aussage zu, dass in der Physik eine absolute Wahrheit gefunden werden kann. Nach Trautwein und Lüdtke (2007) zeigen Schüler in Bezug auf die Themen, die im Unterricht intensiver behandelt wurden, naivere epistemologische Überzeugungen als zu Themen, die weniger intensiv behandelt wurden.

Einen weiteren, ähnlichen Befund liefern auch Maggioni, Alexander und van der Sledright (2004). Aufbauend auf dem Reflective-Judgement-Modell von King et al. (1993) wurden die epistemologischen Überzeugungen von 72 Lehrern, die an einem Unterrichtsprogramm zur

amerikanischen Geschichte teilnahmen, vor der Durchführung des Programms erhoben. Die Erfassung der epistemologischen Überzeugungen erfolgte mithilfe des aus 21 Items bestehenden Fragebogens *Beliefs about Learning and Teaching of History Questionnaire* (BLTHQ) mit einer 6-stufigen Ratingskala. Dieser Fragebogen wurde einer explorativen Faktorenanalyse unterzogen und mittels der Faktoren wurden die Personen als epistemologische Relativisten und kriterienorientierte Denker eingestuft, wobei die kriterienorientierten Denker als epistemologisch fortgeschrittener gelten. Beobachtet wurde die Veränderung der epistemologischen Überzeugung nach Durchführung des Programms, indem die Verschiebung der Personen in andere epistemologische Gruppen erfasst wurde. Die Veränderung der Gruppenzuordnung zeigt Abbildung 26.

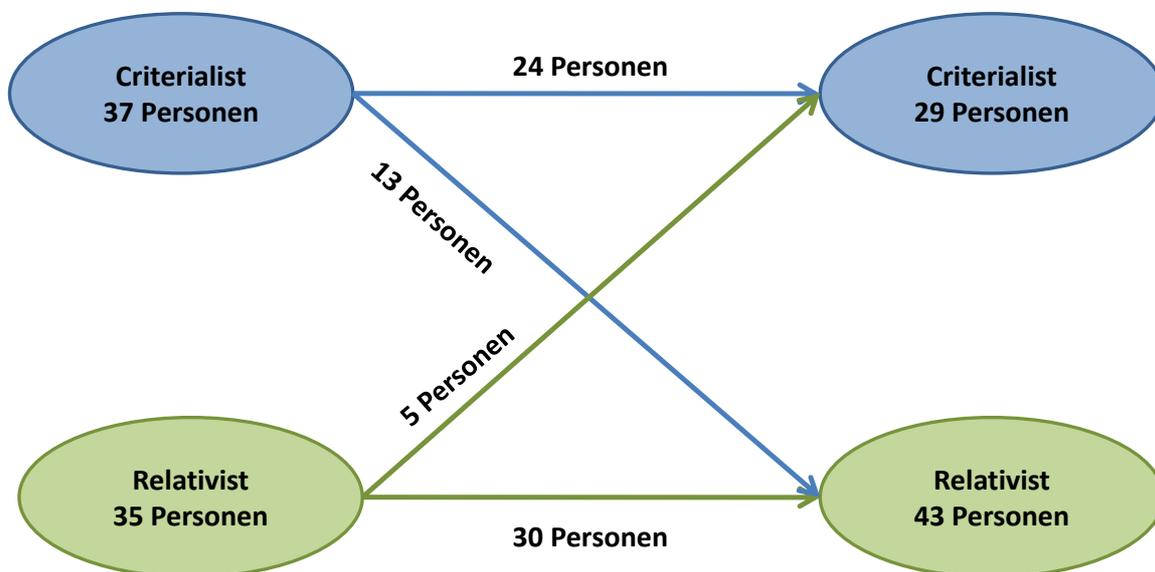


Abbildung 26: Ergebnisse aus Maggioni et al. (2004)

Während also jeweils der Großteil der Teilnehmer in der ersten und zweiten Erhebung in der gleichen Gruppe blieb, erfolgte bei einigen Versuchspersonen ein Gruppenwechsel, bei dem der größere Anteil zu der epistemologisch naiveren Gruppe wechselte. Die vorgestellten theoretischen Überlegungen und empirischen Ergebnisse lassen daher weiteren Zweifel an der üblichen normativen Sichtweise über die Entwicklung epistemologischer Überzeugungen von naiv zu fortgeschritten aufkommen. Mögliche Erklärungen sind noch nicht umfassend ausgearbeitet (Kienhues et al., 2008, Maggioni et al., 2004). Eine Erklärung, die insbesondere mit den Befunden von Köller et al. (2000) übereinstimmt, könnte mit der Entwicklung von Expertise in einer Domäne zusammenhängen. Beschäftigt sich eine Person eine längere Zeit mit einem bestimmten Thema, so werden bestimmte Sachverhalte nicht mehr aktiv reflektiert und als selbstverständlich genommen, sodass in der Folge die epistemologischen Überzeugungen

der Person als naiv erscheinen. Allerdings scheint eine solche Erklärung nicht für eher kurzzeitige Untersuchungen geeignet, wie z. B. der experimentell angelegten Studie von Kienhues et al. (2008), da die Entwicklung von Expertise ein langfristiger Prozess darstellt.

An dieser Stelle ist auch ein Befund relevant, der die Domänenabhängigkeit epistemologischer Überzeugungen betrifft und schon in der Einführung aufgegriffen wurde (vgl. Kapitel 1.5.1, Royce, 1978).⁵⁸ Es konnte weiterhin gezeigt werden, dass in unterschiedlichen akademischen Disziplinen unterschiedliche epistemologische Überzeugungen vorherrschen. So wird z. B. in der Teilchenphysik dem Experiment als Methode der Erkenntnisgewinnung der Vorrang eingeräumt. In der Molekularbiologie hingegen werden zwar ebenfalls Experimente durchgeführt, aber der Forschende tritt als Autorität in den Vordergrund (Knorr Cetina, 1999). Smyth (2001) verglich Lehrbücher verschiedener Disziplinen. Sie fand, dass Lehrbücher aus den naturwissenschaftlichen und mathematischen Disziplinen Fakten ohne Referenz auf eine Quelle aufführen, während psychologische Lehrbücher beim Nennen von Fakten immer auf eine Quelle verweisen. Überdies fand Smyth, dass in Lehrbüchern der Psychologie psychologische Fakten immer mit Referenz präsentiert werden, während naturwissenschaftliche oder mathematische Fakten nur genannt werden. Smyth unterscheidet nach diesem Befund *certain-ty* und *uncertainty sciences*. In den *uncertainty sciences* wie der Psychologie wird nach Smyth nicht die Sicherheit der Aussagen für sich angenommen, sondern es müssen zur Rechtfertigung der Aussagen immer die entsprechenden Evidenzen angegeben werden. Es finden sich also auch hier Unterschiede in den Epistemologien der jeweiligen Domänen. Für unterschiedliche epistemologische Sichtweisen zwischen verschiedenen Domänen sprechen aber auch direkte Befunde der Untersuchungen zur Domänenspezifität epistemologischer Überzeugungen. So konnten z. B. Lonka und Linblom-Yläne (1996) Unterschiede in den epistemologischen Überzeugungen von Studierenden der Medizin und Psychologie nachweisen. Sie fanden, dass unter Psychologiestudierenden hauptsächlich konstruktivistische Vorstellungen über Wissen und Lernen existierten, während diese bei Medizinstudenten hauptsächlich auf die Rezeption von Fakten beschränkt sind und Wissen als eine Ansammlung unverbundener Aussagen angesehen wird. Diese Annahme entspricht der Argumentation von Buehl und Alexander (2006),

⁵⁸ An dieser Stelle soll kurz ein Hinweis auf den Zusammenhang der Entwicklung epistemologischer Überzeugungen und deren Domänenabhängigkeit erfolgen. Die Entwicklung epistemologischer Überzeugungen vollzieht sich immer in einem Kontext, der von den jeweiligen Domänen gebildet wird. Im Zusammenhang mit der o. g. Untersuchung Köller, Baumert und Neubrand (2000) wäre der Kontext der Physikkurs der gymnasialen Oberstufe. Im Kontext eines Studiums wäre dies das Fach, das studiert wird. Insofern sind die Entwicklung epistemologischer Überzeugungen und die Domäne miteinander verwoben.

nach der epistemologischen Überzeugungen durch den soziokulturellen Kontext, in dem sich ein Individuum bewegt, geprägt werden.

Die Befunde zur Ausbildung unterschiedlicher epistemologischer Überzeugungen entsprechen überdies den *Ways of Knowing* von Royce (1978, vgl. Kapitel 1.5.1). Allerdings wird an dieser Stelle nicht die systematische Abhängigkeit epistemologischer Überzeugungen von dem Strukturierungsgrad der jeweiligen Domäne berücksichtigt, wie dies von Royce bei der Entwicklung seines Modells getan wurde. Wichtig an dieser Stelle ist, dass die Befunde eine Abhängigkeit von der Domäne, in welche eine Person im Laufe ihrer akademischen Entwicklung sozialisiert wurde, nahelegen. Infolge dieser Abhängigkeit werden bestimmte, domänenspezifische Kriterien zur Rechtfertigung von Aussagen von einem Individuum präferiert werden, die überdies der üblichen normativen Vorstellung über die Entwicklung von epistemologischen Überzeugungen widersprechen.

6.3.3 Methodische Folgerungen der Kontextabhängigkeit

Die vorgebrachten Einwände gegen eine normative Sichtweise greifen aber noch weiter. So wird in der normativen Sichtweise auch davon ausgegangen, dass zwischen epistemologischen Überzeugungen und anderen Konstrukten, wie z. B. dem Textverstehen, dem Lernerfolg oder auch Lernstrategien, ein positiver Zusammenhang besteht. Diese Position wird aber unter anderem durch die ermittelten Befunde zum Zusammenhang zwischen epistemologischen Überzeugungen und Lernstrategien infrage gestellt. Aus empirischer Sicht zeigt sich zwar ein Zusammenhang epistemologischer Überzeugungen mit Lernstrategien und somit liegt auch eine gewisse pädagogische Relevanz epistemologischer Überzeugungen auf der Hand, allerdings erscheint nach den vorgestellten Befunden der Zusammenhang nur schwach zu sein.

Die Kontextabhängigkeit epistemologischer Überzeugungen liefert hierfür eine mögliche Erklärung. Wenn die Antwort auf die Items abhängig vom jeweiligen Kontext ist, können keine allgemeinen Vorhersagen getroffen werden (Leach et al., 2000). Dies führt aber dazu, dass epistemologische Überzeugungen, wie sie durch Fragebogenitems erfasst werden, zu unspezifisch sind, um konkrete (selbst berichtete) Lernhandlungen vorherzusagen. Diese Sichtweise entspricht einer Problematik, die aus der Einstellungsforschung bekannt ist. Einen engen Zusammenhang zwischen Einstellung und Verhalten gibt es nur dann, wenn die Einstellung und das Verhalten einen gleichen Spezifikationsgrad aufweisen (Ajzen & Fishbein, 1977). Wenn

also nach allgemeinen Einstellungen gefragt wird, kann dadurch kein spezielles Verhalten vorhergesagt werden. Ähnlich ist es bei der Erfassung epistemologischer Überzeugungen. Werden diese in einer dekontextualisierten Weise – wie z. B. durch Fragebögen – erfasst, kann kein Lernverhalten, das immer mit einer bestimmten Situation verbunden ist, vorhergesagt werden. Somit sind bei der Regression von Lernstrategien auf epistemologische Überzeugungen nur geringe Zusammenhänge zu erwarten, wie es die durchgeführten Regressionsanalysen der EPI-Faktoren auf die Lernstrategien, die mithilfe des LIST erfasst wurden, auch ergeben haben.

Neben den vorgestellten empirischen Befunden aus korrelativen und (quasi-)experimentellen Untersuchungen kommen noch Befunde aus den Analysen der Fragebogen zu tragen. Hierbei handelt es sich um psychometrische Befunde, welche die häufig zu beobachtenden niedrigen Reliabilitäten der Skalen zur Erfassung epistemologischer Überzeugungen betreffen, wie sie auch in der vorliegenden Arbeit bei der Analyse des EBQ sowie des EPI auftraten. Dabei müssen zwei Arten von Reliabilität unterschieden werden: einmal die Retestreliabilität und zum anderen Maße der internen Konsistenz.

Schommer (1993b) berichtet von einer Retestreliabilität von .70 bei einer Zeitspanne zwischen erster und zweiter Messung von acht Wochen. Nach Asendorpf (1999) ist im Traitansatz aber die Zeitstabilität der erhobenen Eigenschaften notwendig. Diese Zeitstabilität wird durch die Retestreliabilität ausgedrückt, die durch die Korrelation der Messwerte von erstem und zweitem Zeitpunkt ausgedrückt und für in der Regel eine Mindesthöhe von .80 gefordert wird. Aus dieser geringen Retestreliabilität ergibt sich eine Folgerung, die aus der Veränderungsmessung in der Prozessdiagnostik als Reliabilitäts-Validitäts-Dilemma bekannt ist (Fisseni, 2004). Wenn die Messungen vom ersten und zweiten Zeitpunkt nicht hoch korrelieren, bedeutet dies, dass nur die Unterschiede zwischen den Messungen reliabel erfasst werden und dass sich die erfassten Merkmale zwischen den Messungen verändert haben. Die niedrige Korrelation bei der wiederholten Erhebung epistemologischer Überzeugungen deutet also darauf hin, dass beim zweiten Messzeitpunkt nicht dieselben Konstrukte erfasst wurden wie bei der Erhebung am ersten Messzeitpunkt.

Als Erklärung für diese niedrigen Reliabilitäten gibt Schommer (1993b) an, dass sich die Versuchspersonen in einem Stadium des epistemologischen Übergangs befinden. Diese Idee von Schommer deutet an, dass sich das erhobene Konstrukt zwischen den beiden Messzeitpunkten

verändert haben könnte. Die niedrige Korrelation der beiden Messzeitpunkte liefert empirische Unterstützung für diese Annahme.

Neben der Retestreliabilität liefern aber auch noch die Kennwerte der internen Konsistenz, vornehmlich Cronbachs α bzw. die Skalenreliabilität ρ , Hinweise auf eine mögliche Situationsabhängigkeit epistemologischer Überzeugungen. Die Reliabilitätswerte der Skalen, die in verschiedenen Studien abgeleitet wurden, liegen meist im unteren bis mittleren Bereich. Clarebout et al. (2001) geben in ihrer zusammenfassenden Übersicht zum EBQ-Fragebogen interne Konsistenzen zwischen .45 und .78 für die jeweils entwickelten Skalen an. Clarebout et al. (2001) erhielten bei ihrer Untersuchung mit einer niederländischen Version des EBQ-Fragebogens in der ersten Studie, die zu zwei Skalen führte, interne Konsistenzen von .68 und .71 und in der zweiten Studie, die zu drei Skalen führte, jeweils .59, .53 und .31. Aber nicht nur die aus dem EBQ-Fragebogen abgeleiteten Skalen weisen in der Regel niedrige interne Konsistenzen auf. Auch Schraw et al. (2002) erhielten bei der Entwicklung des *Epistemological Belief Inventory* nur Skalen, deren interne Konsistenzen zwischen .58 und .68 lagen. Wood und Kardash (2002) fanden bei ihrer Analyse des auf Schommer (1990) sowie Jehng (1991) basierenden Fragebogens interne Konsistenzen der gebildeten Skalen zwischen .54 und .74. Auch in den hier durchgeführten Untersuchungen zeigten sich sowohl für die Skalen des EBQ als auch für den EPI niedrige interne Konsistenzen.

Wie in Kapitel 1.3.3 schon angedeutet, diskutieren Moschner und Gruber (2005b) die Konzeption epistemologische Überzeugungen als States und nicht als Traits. Ein State ist ein zeitlich nicht stabiles und von der Situation abhängiges Merkmal (Kelava & Schermelleh-Engel, 2008). Diese Zusammenhänge von States und Traits werden in der Latent-State-Trait-Theorie (LST) beschrieben (Steyer et al., 1992, Steyer et al., 1999). Mithilfe der LST-Theorie lässt sich die Varianz der beobachteten Variablen in Anteile zerlegen, welche die Stabilität und welche die Variabilität angeben. Die Stabilität drückt dabei den Trait-Anteil aus, die Variabilität drückt den State-Anteil der Variablen aus.

Diese Aufteilung der Varianz in Stabilität und Variabilität kann die niedrigen internen Konsistenzen erklären, die sich bei der Analyse von Skalen zur Erfassung epistemologischer Überzeugungen finden. Die interne Konsistenz gibt, wie in Kapitel 3.4.3 aufgeführt, eine Schätzung der Reliabilität einer Skala an. Ein Messinstrument gilt als perfekt reliabel, wenn die damit erhaltenen Testwerte frei von zufälligen Messfehlern sind und umso weniger relia-

bel, je größer die Einflüsse von zufälligen Messfehlern sind (Schermelleh-Engel & Wagner, 2008). In der klassischen Testtheorie wird Reliabilität als der Quotient der Varianz der wahren Werte und der Varianz der beobachteten Werte berechnet. Die wahren Werte gehen dabei auf den Trait zurück, d. h. je mehr Varianz durch die Traitvariablen aufgeklärt wird, desto höhere Reliabilitäten werden erzielt. Niedrige Reliabilitäten bedeuten in diesem Fall also, dass ein großer Anteil der Varianz der beobachteten Werte durch Fehler oder sonstige, nicht berücksichtigte systematische Einflüsse verursacht wird. Diese nicht berücksichtigten Einflüsse können die Erfordernisse der Situation darstellen, die im Rahmen der klassischen Testtheorie unberücksichtigt bleiben und daher dem Messfehler zugerechnet werden, was zu niedrigen Reliabilitäten führt.

In der LST-Theorie hingegen wird davon ausgegangen, dass eine Messung nicht in einem situativen Vakuum stattfindet, sondern dass eine Skala neben den interindividuellen Unterschieden des Traits auch immer systematische Effekte der Situation misst. Daraus folgt aber, dass interindividuelle Unterschiede immer in einem State erhoben werden, wobei diese sich aus den interindividuellen Unterschieden der Traitvariablen, interindividuellen Unterschieden der Statevariablen und Unterschieden in der Situation zusammensetzen.

Im Rahmen der LST-Theorie wird nun Varianzzerlegung der klassischen Testtheorie in Varianz der wahren Werte und Fehlervarianz erweitert. In der LST-Theorie wird die Varianz aufgeteilt in die Varianz der Traits, in die Varianz der States sowie in Messfehlervarianz. Dies hat zur Folge, dass neben dem Anteil der Traits auch der Anteil des States berücksichtigt wird. Eine Beschreibung epistemologischer Überzeugungen mithilfe der LST-Theorie gestattet es also, zu bestimmen, wie hoch der Anteil der Disposition und wie hoch der Anteil der Situation ist. Damit wäre es möglich, den Anteil der Situation empirisch zu bestimmen und die oben genannten ausgeführten situativen Bedingungen epistemologischer Überzeugungen empirisch zu überprüfen.

Die LST-Theorie bietet einen zusätzlichen methodischen Vorteil. So ist es im Rahmen der LST-Theorie nicht erforderlich, zur Erfassung von States und Traits unterschiedliche Instruktionen zum Bearbeiten der Items bzw. unterschiedliche formulierte Items zu nutzen, wie dies z. B. im operationalen Ansatz von Spielberger (1972) der Fall ist. Nach Spielberger erfolgt die Erfassung von States und Traits mithilfe unterschiedlicher Instruktionen. So wird zur Erfassung eines States beispielsweise die Formulierung „Geben Sie bitte an, wie Sie sich jetzt

fühlen“, verwendet, während zur Erfassung eines Traits die Formulierung „Geben Sie bitte an, wie Sie sich im Allgemeinen fühlen“, benutzt wird. Der Ansatz, States und Traits ausgehend von einer unterschiedlichen Instruktion zu erfassen, besitzt aber einen Nachteil. Bei einer solchen Erfassung ist die Beziehung von States und Traits, also die Aufteilung der Varianz und ggf. die Zusammenhänge von States und Traits, nicht zu bestimmen (Steyer et al., 1992). Die Modellierung von epistemologischen Überzeugungen mithilfe der LST-Theorie bietet somit eine Reihe von Vorteilen. Die Idee von Moschner und Gruber (2005b), epistemologische Überzeugungen mithilfe der State-Trait-Unterscheidung zu konzeptionalisieren, bleibt allerdings solange unvollständig, bis diese Unterscheidung auf eine methodische Grundlage gestellt wird, die es gestattet, ihre Vor- und Nachteile abzuwägen sowie eine empirische Überprüfung durchzuführen.

Der Darstellung epistemologischer Überzeugungen mithilfe der LST-Theorie soll noch ein Aspekt epistemologischer Überzeugungen vorangestellt werden, der zuvor bei der Diskussion des Rechtfertigungskriteriums von Rozendaal et al. (2001) angesprochen wurde und ebenfalls mit der normativen Sichtweise epistemologischer Überzeugungen zusammenhängt, und zwar ist dies der soziale Aspekt epistemologischer Überzeugungen.

6.3.4 Soziale Aspekte epistemologischer Überzeugungen

Ein weiterer Aspekt der normativen Sichtweise betrifft die epistemologischen Überzeugungen zur Quelle des Wissens. In der normativen Sichtweise werden in Bezug auf die Quelle des Wissens denjenigen Personen fortgeschrittene epistemologische Überzeugungen zugesprochen, die Wissen als internal ansehen, d. h. jedes Wissen durch eine eigene kognitive Verarbeitung rechtfertigen wollen. Der Bezug auf eine externe Wissensquelle bzw. auf Autoritäten als Quelle des Wissens wird als naiv betrachtet.

Diese normative Sichtweise epistemologischer Überzeugungen, dass Wissen internal erworben wird, ist aber praktisch nicht haltbar. Zwar ist die Vorstellung, dass Wissen am besten durch eigene Erfahrung erworben wird, intuitiv eingängig. Das meiste Wissen wird allerdings durch andere Personen vermittelt. Hinzu kommt, dass eine Person in den wenigsten Bereichen Expertise entwickelt und ausreichend Vorwissen besitzt, um alle Aussagen selbst beurteilen zu können. In den meisten Bereichen gelangt eine Person nicht über den Status eines Laien hinaus, da kein entsprechendes themenbezogenes Vorwissen und ontologisches Wissen über einen Bereich aufgebaut werden kann (Bromme et al., 2008). Bromme, Kienhues und Porsch

(2010) führen an, dass das meiste Wissen in der Ontogenese eines Menschen nicht durch direkte Erfahrung erworben wird. Zu Beginn in der Kindheit wird Wissen von den Eltern, dann von den Lehrern und später den Dozenten übernommen. Später im Erwachsenenalter, wenn Wissen selbstständig erworben werden muss, spielen auch Wissensquellen wie Bücher, Zeitschriften, Fernsehen und das Internet eine wesentliche Rolle bei der Vermittlung von Wissen. Wissen wird in einer Gesellschaft also durch viele verschiedene Quellen vermittelt. Verschiedene Quellen legen auch nahe, dass es für unterschiedliches Wissen auch jeweils unterschiedliche Wissensquellen bzw. Experten für den jeweiligen Bereich geben kann. Wissen ist also über verschiedene Wissensquellen verteilt, wobei sich Wissensquelle in diesem Zusammenhang auf das Medium beziehen kann, in dem das Wissen gespeichert ist oder auf die Experten, welche Wissen in einer Domäne besitzen.

Diese soziale Verteilung von Wissen wird als Paradigma des *distribuierten bzw. verteilten Wissens* bezeichnet (Hutchins, 1995, Salomon, 1997). Im Paradigma verteilten Wissens wird davon ausgegangen, dass Wissen bzw. kognitive Prozesse innerhalb der Mitglieder einer sozialen Gemeinschaft verteilt sein kann bzw. können. Für die Diskussion über epistemologische Überzeugungen sind insbesondere die soziale Verteilung von Wissen unter Mitgliedern einer Gemeinschaft sowie die sich daraus ergebende Konsequenz von Bedeutung, dass Wissen zwischen den Mitgliedern einer Gesellschaft ausgetauscht werden kann. Ebenso wichtig ist, dass einige Mitglieder einer Gesellschaft Wissen in irgendeiner Form in externen Quellen festhalten (z. B. in Form von Büchern, Zeitschriften, Internet etc.) und andere Mitglieder der Gesellschaft darauf zugreifen können. Es sind nun entsprechende Mechanismen erforderlich, um die Rechtfertigung der Aussagen von anderen Gesellschaftsmitgliedern oder aus externen Quellen zu prüfen. Somit sind kognitive Mechanismen erforderlich, die es gestatten, die Glaubwürdigkeit einer Wissensquelle einzuschätzen. Dies wird dadurch verkompliziert, dass ein entsprechendes Vorwissen erforderlich ist, welches aber nicht bei allen Personen als gegeben angesehen werden kann.

Hinzu kommt die in heutigen Gesellschaften anzutreffende Arbeitsteilung, die zu einer Abhängigkeit von Experten eines entsprechenden Wissensbereichs und zu einer ungleichen Verteilung des Wissens in einer Gesellschaft führt. Diese Arbeitsteilung ist auch in der modernen Wissenschaft zu finden (Knorr Cetina, 1999). Dadurch ergibt sich die Notwendigkeit, nicht einzelne Aussagen zu beurteilen und Wissen durch eigene Beschäftigung mit bestimmten Inhalten zu erwerben, sondern es muss die Verteilung des Wissens in einer Gesellschaft er-

kannt werden. Weiterhin müssen Kriterien zur Bestimmung von Experten eines Wissensbereichs bzw. zur Bestimmung von geeigneten Wissensquellen existieren (Baron, 1993). Auch ein Studium ist überdies so gegliedert, dass nur bestimmtes Wissen durch eigene Erfahrung erworben wird, insbesondere bei der Vermittlung der Grundlagen eines Fachs wird vielfach das Lernen von Fakten in den Vordergrund gestellt. Würde jeder Student das schon vorhandene Wissen durch eigene Erfahrungen selbst zu entdecken versuchen, wäre ein zügiges Studium unmöglich. Ein Studium kann als Trainingsprogramm betrachtet werden, in dem bestimmte Fertigkeiten in einem akzeptablen Zeitrahmen vermittelt werden (Schulz, 1993).

Nach Bromme et al. (2010) wurde diesen Aspekten in der üblichen normativen Sichtweise epistemologischer Überzeugungen keine Rechnung getragen. Im Modell von Hofer und Pintrich (1997) beziehen sich die Dimensionen *Wissensquelle* und *Wissensbegründung* auf diese genannten sozialen Aspekte epistemologischer Überzeugungen. In Bezug auf die Dimension *Wissensquelle* ist die übliche normative Sichtweise, dass Wissen, welches durch den Lernenden selbst konstruiert wurde, besser ist als Wissen, welches vom Lernenden von externen Quellen aufgenommen wurde. In Bezug auf die Dimension *Wissensbegründung* geht die normative Sichtweise davon aus, dass die Lernenden Aussagen selbst wie ein Experte auf dem jeweiligen Gebiet beurteilen und damit die Verteilung von Wissen in einer Gesellschaft umgehen. Dies widerspricht aber der bereits vorgetragenen Argumentation.

Die Verteilung von Wissen macht es also erforderlich, nicht mehr die Aussagen direkt zu beurteilen, sondern zu beurteilen, wie glaubwürdig eine Quelle ist. Bromme et al. (2010) trafen die Unterscheidung zwischen *Beurteilungen aus erster Hand* und *Beurteilungen aus zweiter Hand*. Die Frage „Welche Behauptung ist wahr?“ wird damit zu der Frage „Welche Wissensquelle ist glaubwürdig und relevant?“. Die psychologischen Prozesse, welche der Beantwortung beider Fragen zugrunde liegen, sind unterschiedlich. Eine fortgeschrittene epistemologische Überzeugung muss diese Aspekte berücksichtigen.

Allerdings erfordern solche Beurteilungen aus zweiter Hand auch den Aufbau entsprechenden Vorwissens, und zwar über die Verteilung von Wissen in einer Gesellschaft. In den Wissenschaften wird hier üblicherweise eine disziplinbezogene Unterscheidung getroffen, die allerdings oft nicht ausreicht, da verschiedene Themen von mehreren akademischen Disziplinen behandelt werden.

Die Verteilung von Wissen einschließlich darauf spezialisierter psychologischer Mechanismen wurde bis hier nur theoretisch hergeleitet. Offen bleiben empirische Nachweise dieser Mechanismen. Erwachsene haben eine kognitive Repräsentation der Verteilung von Wissen in der Gesellschaft und können diese Verteilung von Wissen erkennen, sie können zusätzliche Informationen aus verschiedenen Wissensquellen abrufen, entsprechende Experten für eine spezifische Problemstellung erkennen und ebenso die Grenzen des eigenen Wissens einschätzen. Die Kategorisierung entsprechender Experten für einen Wissensbereich orientiert sich an den verschiedenen Disziplinen, so wie diese während der Enkulturation im Erziehungssystem vermittelt werden (Keil, Stein, Webb, Billings, Rozenblit, 2008). Entwicklungspsychologische Befunde sprechen dafür, dass schon im Kindesalter ein Verständnis für die Verteilung von Wissen vorliegt. Lutz und Keil (2002) konnten belegen, dass schon 4- bis 5-jährige Kinder die Verteilung des Wissens in den Bereich Medizin und Autotechnik verstehen können, wobei dieses Verständnis nur auf die beobachtbaren Tätigkeiten beschränkt ist. Danovitch und Keil (2004) zeigten, dass sich im Verlauf der Grundschule das Verständnis der Verteilung des Wissens ändert, zuerst wird anhand der jeweiligen Themen differenziert, wer für welches Thema relevantes Wissen besitzt. Später wird diese Differenzierung anhand der Gliederung von Wissen in Disziplinen vorgenommen (vgl. Bromme et al., 2010 für eine zusammenfassende Darstellung dieser Befunde). Bromme et al. (2010) betrachten auch die Befunde von Kuhn (1991) kritisch, wonach nur ein geringer Teil der untersuchten Personen die höchste Stufe in deren Stufenmodell, in welcher die Verteilung des Wissens in einer Gesellschaft bei der Begründung von Aussagen anerkannt wird. Die Wissensverteilung in Gesellschaften macht es psychologisch unplausibel, dass die Verteilung von Wissen bei der Beurteilung von Aussagen und Wissensquellen nur von einem geringen Teil der Mitglieder einer Gesellschaft beachtet wird. Die Anerkennung der Verteilung von Wissen ist nicht die Ausnahme, sondern stellt den Normalzustand dar.

Abschließend soll der Frage nachgegangen werden, ob dieser soziale Aspekt epistemologischer Überzeugungen in den bisherigen Fragebogen schon bedacht wurde und ob es sich um eigene epistemologische Überzeugungen handelt, die sich auf die Verteilung von Wissen beziehen und die noch nicht in den gängigen Inventaren vorhanden sind. Bislang wurde nur in der Konzeption von Moschner et al. (2005) eine soziale Komponente epistemologischer Überzeugungen berücksichtigt, die sich allerdings auf die Rechtfertigung von Aussagen als ein gemeinsames Produkt einer Scientific Community bezieht. Das kann beispielsweise Wissen betreffen, das in von der Scientific Community anerkannten Literatur publiziert wurde.

Diese Dimension bezieht sich somit auf die Grundlagen, die zur Rechtfertigung einer Aussage benutzt werden, in diesem Fall ist dies der Grad, in dem eine Aussage von anderen Personen in der Scientific Community akzeptiert wird.

Bromme et al. (2010) sprechen sich dafür aus, die Verteilung von Wissen in die Dimension *Wissensstruktur* zu integrieren. Dies stellt aber wie zuvor aufgeführt keine Komponente epistemologischer Überzeugungen dar, da sich die Struktur des Wissens nicht auf das Rechtfertigungskriterium bezieht. Allerdings steht die Struktur des Wissens in einer Beziehung zu den epistemologischen Überzeugungen einer Person. Die Frage nach der Beziehung von strukturellen Aspekten des Wissens gliedert sich in die Frage ein, die bei den beiden Dimensionen *Geschlechtsspezifische Wissenszugänge* und *Kulturspezifische Wissenszugänge* gestellt wurde. Beide Dimensionen stellen nach dem Rechtfertigungskriterium ebenfalls keine epistemologischen Überzeugungen dar, allerdings handelt es sich auch hier um Konstrukte, die in einer Beziehung zu epistemologischen Überzeugungen stehen und diese beeinflussen können.

Es kann aber auch argumentiert werden, dass der soziale Charakter epistemologischer Überzeugungen sich nicht in einer weiteren Dimension ausdrückt, sondern in der Art und Weise der Nutzung des Rechtfertigungskriteriums. Wie bereits aufgeführt wurde, wird bei Wissen, das verteilt ist und bei dem eine Wissensvermittlung durch Autoritäten stattfindet, eine externe Rechtfertigung benötigt.

An dieser Stelle kann als Fazit der letzten drei Kapitel zusammengefasst werden, dass fortgeschrittene epistemologische Überzeugungen als kontextabhängige Beurteilungen von Aussagen aufgefasst werden können und dass für die Beurteilungen von Aussagen aus erster und zweiter Hand die Verteilung von Wissen in einer Gesellschaft berücksichtigt werden muss. Die Kontextabhängigkeit und die Verteilung von Wissen bedingen interkontextuelle Varianz. Epistemologische Überzeugungen hängen nicht nur von der Person, sondern auch von der Situation sowie der Interaktion von Person und Situation ab. Deswegen führt auch dies dazu, dass epistemologische Überzeugungen als States und nicht als Traits formuliert werden sollten (vgl. den Vorschlag von Muschner und Gruber (2005b), der in Kapitel 5.2 dargestellt wurde).

6.4 Epistemologische Überzeugungen als Schemata

In den vorherigen Kapiteln der Diskussion wurden Fragen besprochen, die sich während der Auswertung des EPI- und EBQ-Fragebogens aufgetan haben. Es wurde zunächst ein Kriterium aufgestellt, das es erlaubt, epistemologische Überzeugungen von anderen Konstrukten abzugrenzen. Anschließend wurden verschiedene Probleme erörtert, die sich bei der Erhebung epistemologischer Überzeugungen durch Fragebogen ergeben, die sich unter dem Stichwort Validität zusammenfassen lassen. Darüber hinaus wurden verschiedene Probleme betrachtet, die sich aus der normativen Sichtweise ergeben, die der Erfassung epistemologischer Überzeugungen zugrunde liegt. Dies waren die Kontextabhängigkeit epistemologischer Überzeugungen sowie deren sozialer Charakter. In diesem Kapitel soll eine kognitive Architektur vorgestellt werden, die es erlaubt, die dargestellten Aspekte abzubilden und die im Folgenden als Ansatzpunkt zu einem eigenen Ansatz epistemologischer Überzeugungen genutzt werden kann.

Epistemologische Überzeugungen lassen sich mithilfe von Schemata modellieren (vgl. Bartlett, 1932, Bless & Schwarz, 2002, Kopp & Mandl, 2006 zum Schemabegriff und die Möglichkeiten, die dieser für Modellierung kognitiver Strukturen bietet). Schemata lassen sich als allgemeine Denkgeln verstehen, die sich im Fall epistemologischer Überzeugungen auf den Gegenstand Wissen beziehen. Schemata organisieren die vergangenen Erfahrungen und Handlungen und bilden Repräsentationen dieser Handlungen und Erfahrungen. Das Wirken der Schemata kann in jedem Individuum, das an seine Umwelt angepasst ist, angenommen werden (Bartlett, 1932). Schemata beschreiben Merkmale und die Zusammenhänge dieser Merkmale innerhalb eines Gegenstandsbereichs und repräsentieren Wissen über den Gegenstandsbereich auf einem höheren Abstraktionsniveau. In diesem Sinn beziehen sich die Schemata, die epistemologische Überzeugungen beinhalten, auf Wissen über den Gegenstandsbereich Wissen. Die für ein Schema wichtigen Merkmale werden auf diesem Abstraktionsniveau durch Variablen repräsentiert, die bei Aktivierung eines Schemas mit Werten besetzt werden, was auch als Instantiierung eines Schema bezeichnet wird. Die Werte dieser Variablen können allerdings nicht beliebig variieren, sondern sind bestimmten Begrenzungen ausgesetzt, die dafür sorgen, dass nur sinnvolle Werte der Variablen durch ein Schema verarbeitet werden. Werden bei der Aktivierung eines Schemas diese Variablen nicht durch Werte besetzt, wird ein Standardwert aufgerufen. Schemata können weiterhin hierarchisch organisiert sein, sodass auch einer ggf. vorhandenen Hierarchie in einem Gegenstandsbereich Rechnung getragen werden kann. Zudem können mehrere Schemata zu einem semantischen Netzwerk

verbunden werden, sodass die Aktivierung eines Schemas automatisch die Aktivierung von verbundenen Schemata in einem Prozess der Aktivationsausbreitung verursacht (vgl. Loftus, 1975). Häufig abgerufene Schemata können auf diese Weise auch geprimt werden (Bless & Schwarz, 2002). Schemata haben neben dem statischen Aspekt der Abbildung eines Gegenstandsbereichs auch eine Prozesskomponente (Rumelhart, 1980, Kopp & Mandl, 2005). Die Prozesskomponente ermöglicht den aktiven Erwerb von Schemata und deren Verwendung in unterschiedlichen Situationen. Bei der Anwendung von Schemata unterscheidet man zwischen Top-down- und Bottom-up-Informationsverarbeitung. Bei einem Top-down-Prozess ist ein Schema bereits aktiviert, das im Folgenden die Verarbeitung der aufgenommenen Information steuert. Bei einem Bottom-up-Prozess wird ein Schema entsprechend den aufgenommenen Informationen aktiviert.

Überdies bieten Schemata aber noch weitere Vorteile bei der Modellierung epistemologischer Überzeugungen. So sind Schemata seit ihrer ersten Verwendung von Bartlett (1932) im Rahmen gedächtnispsychologischer Untersuchungen nicht nur als passive Verarbeitungsmechanismen konzipiert, sondern als Mechanismen, die eine aktive Verarbeitung der aufgenommenen bzw. abgerufenen Informationen vornehmen und dadurch zur Konstruktion von Wissen fähig sind. Diese aktive Verarbeitung ist aber auch von dem Vorwissen abhängig, das eine Person in die jeweilige Situation einbringt (Bless & Schwarz, 2002). Somit kann der Bedeutung des Vorwissens bei der Rechtfertigung von Aussagen Rechnung getragen werden. Überdies sind Schemata flexibel, sodass sie durch Entwicklungsprozesse verändert werden können. Einerseits kann ein Schema durch Akkommodation an das Vorliegen neuer, abweichender Informationen an diese angepasst werden, andererseits können neue, abweichende Informationen durch Assimilation in ein Schema integriert werden (vgl. Piaget, 1955). Durch zusätzlich Erfahrungen in einem Gegenstandsbereich werden Schemata in der Regel besser organisiert, wobei diese Veränderungen aber erfahrungsgemäß nicht plötzlich, sondern in einem länger dauernden Prozess stattfinden (Bless & Schwarz, 2002). Auf diese Weise können geringere Abweichungen von den im Schema abgelegten Regeln in dieses assimiliert werden, wohingegen Abweichungen, welche das Schema dysfunktional werden lassen, zu einer Akkommodation führen, bei welcher das Schema an sich verändert wird. Wie die Veränderungen von Schemata ablaufen, kann dann z. B. mithilfe der Veränderungsmodelle von Weber und Crocker (1983) beschrieben werden.

Wird die Rechtfertigung einer Aussage verlangt, wird je nach Kontext ein passendes Schema aktiviert, d. h. es findet eine Bottom-up-Verarbeitung statt. Eine solche kognitive Architektur gestattet den flexiblen Einsatz epistemologischer Überzeugungen im Sinne der genannten kontextsensitiven Beurteilung und Rechtfertigung von Aussagen (Bromme et al., 2008). Schemata bieten überdies den Vorteil, dass auch keine Unitaritätsannahme im Sinn von Hammer und Elby (2002) notwendig ist. Im Gegensatz zu dem von Hammer und Elby vorgeschlagenen Modell epistemologischer Ressourcen bieten Schemata weiterhin den Vorteil, dass mit ihnen ein sparsames kognitives Modell epistemologischer Überzeugungen aufgestellt werden kann, ohne dass eine Vielzahl von Ressourcen für unterschiedliche Situationen postuliert werden muss.

Dass epistemologische Überzeugungen als Denkregeln verstanden werden können, konnte empirisch in verschiedenen Studien nachgewiesen werden. Hierzu wurde in Anlehnung an Bartlett (1932) das Wirken der Schemata bei der Beurteilung der Rechtfertigung von Aussagen nachgewiesen. Hierfür sind Methoden erforderlich, welche es gestatten, die in einem Individuum ablaufenden kognitiven Prozesse abzubilden. Ein Beispiel für eine solche Studie ist die in Kapitel 1.5.3 aufgeführte Untersuchung von Hofer (2004) zur Beurteilung von Aussagen beim Suchen von Informationen im Internet. Zu ähnlichen Ergebnissen kommen auch Mason und Boldrin (2008), die mithilfe von lautem Denken epistemologische Überzeugungen ebenfalls bei einer Informationssuche im Internet erhoben. Es zeigten sich für Studenten und Schüler Hinweise auf alle vier Dimensionen des Modells von Hofer und Pintrich (1997). Überdies zeigte die Studie, dass die epistemologische Beurteilung abhängig vom domänenspezifischen Vorwissen ist. So wiesen Personen mit hohem Vorwissen in einer Domäne die Tendenz auf, Informationen aus dem Internet als unsicher zu bezeichnen, während dieselben Personen in einer anderen Domäne, in der sie wenig Vorwissen besaßen, gänzlich andere Beurteilungen zeigten. Da sowohl bei Hofer (2004) als auch bei Mason und Boldrin (2008) epistemologische Überzeugungen mithilfe der Methode des lauten Denkens erhoben wurden, können auch Informationen über den Prozess gesammelt werden, wie eine Person zu einer epistemologischen Beurteilung einer Aussage gelangt. Die Erhebung von Daten mithilfe des lauten Denkens ermöglicht es, dabei Einblicke in die Art und Weise zu erhalten, wie Versuchspersonen bei der Bearbeitung von Aufgaben zu epistemologischen Schlussfolgerungen gelangen. Die Hinweise auf epistemologische Überzeugungen nach dem Modell von Hofer und Pintrich (1997), die sich bei Auswertung der Protokolle des lauten Denkens ergaben, sind

somit auch erste Hinweise auf die Denkregeln oder Schemata, die bei der Beurteilung von Aussagen herangezogen werden.

Eine weitere und sehr detaillierte Studie zum Nachweis epistemologischer Überzeugungen mithilfe der Methode des lauten Denkens wurde von Ferguson, Braten und Stromso (2012) durchgeführt. Diese Untersuchung ist insofern besonders, als dass nicht nur die Beurteilung von Information verlangt wurde, sondern den Probanden gezielt sich widersprechende Informationen zum Thema Gesundheitsrisiken durch Mobilfunk vorgegeben wurde, zudem wurde in der Untersuchung das domänenspezifische Vorwissen berücksichtigt. Eine weitere Besonderheit der Studie in Bezug auf die Auswertung der Laut-Denken-Protokolle sind die berücksichtigten Dimensionen epistemologischer Überzeugungen, so wurde besonderes Augenmerk auf diejenigen Dimensionen gelegt, welche sich mit der Rechtfertigung von Aussagen befassen. Es wurden die Dimensionen Sicherheit und Einfachheit von Wissen nach Hofer und Pintrich (1997) untersucht. In Bezug auf die Rechtfertigung von Aussagen wurden die Dimensionen Rechtfertigung durch eine Autorität, Persönliche Rechtfertigung und Rechtfertigung durch multiple Quellen (vgl. Greene et al., 2008) unterschieden. Untersucht wurden schließlich die Laut-Denken-Protokolle von 51 Probanden, von denen 43 Frauen und acht Männer waren. Die von den Probanden protokollierten spontanen Äußerungen wurden im Hinblick auf die genannten Aussagen kodiert. Tabelle 121 zeigt den Prozentsatz der Probanden, deren Äußerungen sich auf eine der Dimensionen beziehen, Tabelle 122 zeigt die Anzahl der Äußerungen in jeder Dimension.

Tabelle 121: Prozentualer Anteil der Probanden mit spontanen Äußerungen über epistemologische Überzeugungen (Ferguson et al., 2012)

Dimension	Probanden	
	Anzahl	Prozentsatz
Sicherheit/Einfachheit (S/E)	47	92.2
Rechtfertigung durch Autorität (RA)	27	52.9
Persönliche Rechtfertigung (PR)	16	31.4
Rechtfertigung durch multiple Quellen (MQ)	25	49.0
S/E+RA	5	9.8
S/E+PR	5	9.8
S/E+MQ	3	5.9
S/E+RA+PR	3	5.9
S/E+RA+MQ	14	27.5
S/E+PR+MQ	2	3.9
S/E+RA+PR+MQ	6	11.8

Tabelle 122: Nach Dimensionen aufgeteilte Prozentsätze aller abgegebenen Äußerungen (Ferguson et al., 2012)

Dimension	Äußerungen	
	Anzahl	Prozentsatz
Sicherheit/Einfachheit (S/E)	205	63.9
Rechtfertigung durch Autorität (RA)	39	12.1
Persönliche Rechtfertigung (PR)	28	8.7
Rechtfertigung durch multiple Quellen (MQ)	49	15.3
Gesamtzahl der Äußerungen	321	100

Die Untersuchung von Ferguson et al. (2012) belegt somit sehr ausführlich, dass Probanden sich bei der Beurteilung von gegensätzlichen Aussagen auf epistemologische Überzeugungen stützen.

Nachdem epistemologische Überzeugungen bzw. die Schemata zur Rechtfertigung von Aussagen weitestgehend im akademischen Kontext und in Lehr-Lern-Situationen betrachtet wurden, obwohl sie sich nicht nur auf den akademischen Bereich bzw. auf Lehr-Lern-Situationen beschränken, ist darauf einzugehen, dass und wie sie in alltäglichen Kontexten feststellbar sind, wie es die Untersuchung von Ferguson et al. (2012) zu den Effekten von elektromagnetischer Strahlung exemplarisch zeigt. Einen weiteren alltäglichen Kontext zur Beurteilung der Rechtfertigung von Aussagen bilden medizinische Informationen. Bei medizinischen Informationen kann aufgrund der gesellschaftlichen Verteilung von Wissen angenommen werden, dass die meisten Menschen nicht über das entsprechende Vorwissen verfügen und somit Laien sind. In solchen Kontexten ist insbesondere die Beurteilung der Quellen von Bedeutung, die Behauptungen liefern bzw. stützen. Wittwer, Bromme und Jucks (2004) haben die Beurteilung der Glaubwürdigkeit von medizinischen Informationen in den Medien Internet und Zeitschriften untersucht. Untersucht wurde die Glaubwürdigkeit einer bestimmten Wissensquelle. Bezogen auf epistemologische Überzeugungen geht es hier um die Frage, wie die Rechtfertigung von Aussagen in Bezug auf eine bestimmte Wissensquelle zu beurteilen ist. Die Beurteilung der Rechtfertigung von Aussagen und damit zusammenhängender Aspekte wie die Glaubwürdigkeit von Experten oder Medien sind Handlungen, die alltäglich durchgeführt werden müssen (vgl. Bromme et al., 2010). An dieser Stelle wird deutlich, dass es sich bei epistemologischen Überzeugungen nicht um ein Konzept handelt, das nur im Rahmen von Lehr-Lernvorgängen Anwendung findet. Vielmehr stellen epistemologische Überzeugungen kognitive Ressourcen dar, die in vielen Kontexten angewendet werden. Epistemologische Überzeugungen stellen somit ein Konstrukt dar, das nicht nur im Rahmen pädagogisch-

psychologischer Forschung relevant, sondern für die gesamte Psychologie, die sich mit dem Erwerb von Wissen beschäftigt, wichtig ist. Zu nennen ist hier insbesondere die medienpsychologische Forschung zum Wissenserwerb mit Neuen Medien wie dem Internet oder dem Fernsehen. Aber auch für organisationspsychologische Fragestellungen zum Wissenserwerb in Organisationen stellen epistemologische Überzeugungen wichtige Faktoren dar.

Dass epistemologische Überzeugungen nicht nur in instruktionalen Kontexten eine Rolle spielen, wird im TIDE-Modell von Muis et al. (2006) berücksichtigt. Hier wird zwischen allgemeinen, akademischen und domänenspezifischen epistemologischen Überzeugungen unterschieden. Allgemeine epistemologische Überzeugungen beziehen sich auf nicht akademische Kontexte, wie z. B. bei der Beurteilung der Rechtfertigung von Aussagen im Rahmen der Wissensvermittlung in unterschiedlichen Medien oder in der Interaktion mit anderen Menschen. Akademische epistemologische Überzeugungen beziehen sich auf akademische Kontexte der Wissensvermittlung und sind über Domänen hinweg gleich. Domänenspezifische epistemologische Überzeugungen sind auf eine bestimmte Domäne bezogene epistemologische Überzeugungen und werden hauptsächlich in instruktionalen Kontexten einer Domäne herausgebildet. Bei den allgemeinen akademischen und domänenspezifischen epistemologischen Überzeugungen handelt es sich nicht jeweils um besondere Formen, sondern es werden allgemeine kognitive Ressourcen zur Beurteilung von Aussagen und Wissensquellen entwickelt, die in jeweils dem einem oder anderen Kontext angewendet werden. Das TIDE-Modell vereint somit die Befunde zur Kontext- und Domänenabhängigkeit von epistemologischen Überzeugungen. Die Konzeption von epistemologischen Überzeugungen als Schemata gestattet eine entsprechende Flexibilität in der Modellierung epistemologischer Überzeugungen nach dem TIDE-Modell. Die Darstellung der wesentlichen Merkmale von epistemologischen Überzeugungen als Variablen eines Schemas sowie die potenziell hierarchische Organisation von Schemata bieten die notwendigen Möglichkeiten, um die Taxonomie epistemologischer Überzeugungen nach dem TIDE-Modell abzubilden.

Die Gliederung von allgemeinen, akademischen und domänenspezifischen epistemologischen Überzeugungen ist auch mit der Gliederung von Wissensformen nach De Jong und Ferguson-Hessler (1996) vereinbar. De Jong und Ferguson-Hessler teilen Wissen in die zwei Dimensionen Wissensart und Wissensmerkmal ein. Die Wissensart gibt an, welches Wissen vorliegt, z. B. prozedurales oder konzeptuelles Wissen. Die Wissensmerkmale geben verschiedene Eigenschaften an, welche eine Wissensart aufweist. Unter anderem wird hier der Allgemein-

heitsgrad einer Wissensart beschrieben, der angibt, ob eine Wissensart eher allgemein oder eher domänenspezifisch ist. Neben dem Allgemeinheitsgrad einer Wissensart gibt es aber weitere Parallelen, welche die Wissenstaxonomie von De Jong und Ferguson-Hessler mit der Forschung über epistemologische Überzeugungen teilt. So gibt die innere Struktur an, ob eine Wissensart aus isolierten Wissenseinheiten besteht oder vernetztes Wissen darstellt. Und die hierarchische Struktur gibt an, ob eine Wissensart eher durch eine oberflächliche oder eine tiefe Verarbeitung gekennzeichnet wird. Entsprechende Dimensionen epistemologischer Überzeugungen wären z. B. *Reflexive Natur des Wissens* nach Moschner und Gruber (2005), *Struktur des Wissens* nach Hofer und Pintrich (1997) und *Quick Learning* nach Schommer (1993).⁵⁹ Wie diese Beispiele belegen, sollte die Forschung über epistemologische Überzeugungen und der Psychologie des Wissens immer mit entsprechenden Aspekten der Philosophie integrativ betrachtet werden.

Im Folgenden werden die verschiedenen Befunde in ein gemeinsames Modell integriert. Die Aufstellung dieses Modells orientiert sich an dem in Kapitel 1.5.2 vorgestellten TIDE-Modell von Muis et al. (2006), auf das auch zuvor Bezug genommen wurde. Ausgehend von dem TIDE-Modell wird eine Latent-State-Trait-Theorie epistemologischer Überzeugungen entwickelt.

6.5 Eine Latent-State-Trait-Theorie epistemologischer Überzeugungen

6.5.1 Inhaltliche Darstellung der LST-Theorie epistemologischer Überzeugungen

Im Folgenden wird eine Theorie epistemologischer Überzeugungen anhand der oben diskutierten theoretischen Aspekte entwickelt. Ausgangspunkt für die Entwicklung dieser Theorie ist der in der vorliegenden Arbeit benutzte Traitansatz bzw. die persönlichkeitspsychologische Sichtweise epistemologischer Überzeugungen, der mithilfe eines Stateansatzes zu einer Latent-State-Trait-Theorie epistemologischer Überzeugungen ergänzt werden soll. Im ersten Schritt wird die neue Theorie auf inhaltlicher Ebene als Modifikation des TIDE-Modells von Muis et al. (2006) aufgebaut. Daran anschließend wird die Theorie aus methodischer Sicht mithilfe der Latent-State-Trait-Theorie von Steyer et al. (1992) betrachtet. Die Kombination der psychologischen Theorie mit einem methodischen Latent-State-Trait-Ansatz bietet den

⁵⁹ Angemerkt werden muss, dass die letzten beiden Dimensionen keine epistemologischen Überzeugungen nach dem zuvor eingeführten Abgrenzungskriterium darstellen.

Vorteil, dass die Theorie empirisch prüfbar wird. Die Darstellung ist allerdings nicht als umfassend ausformulierte Theorie anzusehen, sondern als Theorieskizze, die aus den hier diskutierten Arbeiten und empirischen Befunden folgt und die zukünftig weiter ausgearbeitet werden kann.

Epistemologische Überzeugungen stellen kognitive Mechanismen dar, welche die Bewertung der Rechtfertigung von Aussagen auf deren Gültigkeit bzw. die Bewertung einer Wissensquelle, die eine Aussage tätigt, erlauben. Diese kognitiven Mechanismen lassen sich in Form von Schemata modellieren. Epistemologische Überzeugungen können danach unterteilt werden, auf welchen Kontext sie sich beziehen (Bell & Linn, 2002, Muis et al., 2006). *Allgemeine epistemologische Überzeugungen* beziehen sich auf einen alltäglichen Kontext während sich *akademische epistemologische Überzeugungen* auf akademische Kontexte beziehen. Akademische Kontexte beziehen sich auf Umfelder, die in der Schule, der Aus- oder Weiterbildung oder im universitären Bereich auftreten können. Diese akademischen epistemologischen Überzeugungen unterscheiden sich nicht von allgemeinen epistemologischen Überzeugungen, lediglich der Kontext, auf den sie sich beziehen, ist ein anderer. Dabei sind alltägliche und akademische Kontexte nicht strikt voneinander getrennt, sondern können sich gegenseitig beeinflussen. Akademische epistemologische Überzeugungen enthalten jedoch spezielle Rechtfertigungskriterien, die sich auf akademische Inhalte beziehen (Bell & Linn, 2002). *Domänspezifische epistemologische Überzeugungen* spielen in einem speziellen instruktionalen Kontext eine Rolle und sind in diesen eingebettet, wie er sich z. B. bei der Wissensvermittlung in einer einzelnen Domäne ergibt. Domänspezifische epistemologische Überzeugungen setzen ein hohes themenbezogenes und ontologisches Wissen in der jeweiligen Domäne voraus (Bromme et al., 2008).

In Anlehnung an Muis et al. (2006) wird davon ausgegangen, dass sich die epistemologischen Überzeugungen eines Individuums von Geburt an bis zu seinem Tod entwickeln. Weiter wird in Anlehnung an Hammer und Elby (2002) davon ausgegangen, dass Kinder von Geburt an über kognitive Mechanismen zur Beurteilung der Rechtfertigung von Aussagen und Wissensquellen verfügen und dass diese kognitiven Mechanismen die Grundlage für die spätere Entwicklung epistemologischer Überzeugungen bilden. Die Entwicklung epistemologischer Überzeugungen beginnt in der Kindheit und bereits Kleinkinder besitzen die Fähigkeit, die Rechtfertigung und die Quelle von Aussagen zu beurteilen (vgl. Bromme et al., 2010, Burr & Hofer, 2006). Die Entwicklung epistemologischer Überzeugungen ist dabei eingebettet und

wird gleichzeitig beeinflusst durch die jeweilige Kultur bzw. den soziokulturellen Kontext, in der bzw. in dem eine Person aufwächst. Auf diese Weise bilden sich im Laufe der Ontogenese eines Menschen die *allgemeinen epistemologischen Überzeugungen* heraus (vgl. Bell & Linn, 2002).

Mit dem Eintritt in das Bildungssystem beginnen sich epistemologische Überzeugungen zu entwickeln, die sich auf akademische Sachverhalte beziehen (Muis et al., 2006). Aus den allgemeinen epistemologischen Überzeugungen entwickeln sich mit der Zeit die *akademischen epistemologischen Überzeugungen*.

Schreitet eine Person im Erziehungssystem voran (z. B. durch eine weiterführende Schule oder ein Studium), entwickeln sich mit zunehmendem themenbezogenem und ontologischem Wissen in einer Domäne entsprechende *domänenspezifische epistemologische Überzeugungen*. Diese Entwicklung reicht von der Entwicklung eines Novizen, der über kein spezialisiertes Wissen verfügt, bis zu einem Experten, der über spezialisiertes Wissen in einer Domäne verfügt (Royce, 1978).

Werden epistemologische Überzeugungen z. B. per Fragebogen oder anderen Instrumenten erfasst, erfolgt die Messung nicht in einem situationalen Vakuum, vielmehr ist die Person bei der Messung immer in einem allgemeinen bzw. akademischen oder instruktionalen Kontext. Daher befinden sich die Personen bei jeder Messung immer in einer anderen Situation. Dies ist gleichbedeutend mit der Prämisse der Latent-State-Trait-Theorie, dass menschliches Verhalten, Emotion und Kognition in systematischer Weise von der Person, den Charakteristiken der Situation sowie der Interaktion von Person und Situation abhängen. Die Abhängigkeit von der Person spiegelt sich in der individuellen Ausprägung einer interessierenden Variablen wieder, was als interindividuelle Variation aufgefasst wird. Die beiden Faktoren Person und Situation sowie deren Interaktion schaffen einen psychologischen Zustand, der über die Zeit proportional mit der Änderung der Situation variiert (Steyer et al., 1999). Dies wird als intraindividuelle Variation aufgefasst. Im Alltag sind interindividuelle und intraindividuelle Variation somit konfundiert.

Folglich erfasst ein Messinstrument immer interindividuelle Unterschiede und intraindividuelle Variation. Ein Messwert stellt daher eine Zusammenfassung der Ausprägung des Traits, des States sowie der Interaktion von Trait und State dar. Dabei wird unter Trait ein zeitlich stabi-

les, zustands- und situationsunabhängiges Merkmal und unter State ein zeitlich instabiles, zustands- und situationsabhängiges Merkmal verstanden (Kelava & Schermelleh-Engel, 2008). Die Traitkomponenten kommen dadurch zustande, dass die situativen Besonderheiten beim Befragen einer größeren Anzahl von Personen durch Gemeinsamkeiten im Antwortmuster überlagert werden, die sich als interindividuellen Differenzen zeigen (Leach et al., 2000).

Nach der Latent-State-Trait-Theorie muss die Situation, in der sich eine Person befindet, nicht bekannt sein. Die natürliche Variation der Situation von Messzeitpunkt zu Messzeitpunkt sowie die natürlichen Unterschiede zwischen den Personen bei jedem der Messzeitpunkte sind ausreichend. Die LST-Theorie kommt mit der Annahme aus, dass zwischen zwei Messzeitpunkten eine unterschiedliche psychologische Situation vorliegt, die das Verhalten, die Emotionen und die Kognitionen von Personen beeinflusst. Diese unterschiedlichen Situationen müssen überdies objektiv beschrieben sein, damit ihr Einfluss geschätzt werden kann (Steyer et al., 1999).

Neben dem Vorteil der Trennbarkeit von State und Trait bietet die LST-Theorie den Vorteil, dass ihre Anwendung nicht auf Daten beschränkt ist, wie sie z. B. bei der Verwendung von Ratingsskalen resultieren. So ist es beispielsweise auch möglich die LST-Theorie auf EEG Daten anzuwenden (Hagemann, Hewig, Seifert, Naumann & Bartussek, 2005). Insbesondere kann die LST-Theorie auch bei ordinalen Daten angewandt werden (Eid, 1997, Moustaki, 2000) bzw. es können die Methoden der ordinalen Faktorenanalyse (Flora & Curran, 2004) auf ein nach der LST-Theorie spezifizierten Modell angewandt werden. Überdies können auch Mischverteilungsmodelle auf LST-Modelle angewandt werden (Courvoisier, Eid, Nussbeck, 2007). Die LST-Theorie epistemologischer Überzeugungen ist somit nicht nur auf die üblichen Datenbogen aus Fragebögen angewiesen, in den die Zustimmung zu einem Item mittels einer Ratingsskala erhoben wird. Es können z. B. auch Items konstruiert werden, deren Antwortkategorien in geordneter Weise die Ausprägung der Berteilung einer Aussagenrechtfertigung angeben. Die geordneten Antwortkategorien stellen dann ein ordinales LST-Modell dar.

Ein weiterer Vorteil der LST-Theorie ist, dass diese keine Bekanntheit der Situation voraussetzt, es aber offenhält, die Situation, in der sich eine Situation befindet gezielt zu verändern (Kelava & Schermelleh-Engel, 2008). Es können somit auch die Auswirkungen experimenteller Veränderungen auf epistemologische Überzeugungen oder die Auswirkungen von instruktionalen Maßnahmen in einer Domäne untersucht werden. Alles in allem bietet die LST-

Theorie ein flexibles Werkzeug zur Modellierung epistemologischer Überzeugungen. Die methodischen Grundlagen und welche Fragen die LST-Theorie beantworten kann wird im nächsten Kapitel dargestellt.

6.5.2 Methodische Darstellung der LST-Theorie epistemologischer Überzeugungen

Die Latent-State-Trait-Theorie stellt eine Erweiterung der klassischen Testtheorie dar. Der Messwert x_{it} einer Person i zum Zeitpunkt t wird als eine Realisierung einer Zufallsvariablen X_{it} angesehen, welche die intraindividuelle Verteilung der Messwerte x einer Person i zum Zeitpunkt t beschreibt (Kelava & Schermelleh-Engel, 2008). Wie in der klassischen Testtheorie wird der beobachtete Wert einer Person in einem wahren Wert τ und einen Messfehler ε zerlegt, dies lässt sich schreiben als:

$$x = \tau + \varepsilon.$$

In der klassischen Testtheorie gibt der wahre Wert den Anteil des beobachteten Wertes wieder, der auf den zugrunde liegenden Trait zurückzuführen ist und der Messfehler ε – auch Residuum genannt – einen (un-)systematischen Fehleranteil am Gesamtwert x der Messung. Die Reliabilität wird als der Quotient der Varianz der wahren Werte und der Varianz der beobachteten Werte berechnet (Lord & Novick, 1968, vgl. Kapitel 3.4.1). In der klassischen Testtheorie hängt der beobachtete Wert x also nur von dem Trait als Personenvariable ab. Die Annahme der LST-Theorie dagegen ist, dass der wahre Wert nicht nur von der Person, sondern auch von der Situation sowie der Interaktion von Situation und Person bedingt ist. Der wahre Wert wird deswegen hier als *latenter State-Wert* τ bezeichnet und in zwei Komponenten zerlegt, das sind der *latente Trait-Wert* ξ und das *latente State-Residuum* ζ (Kelava & Schermelleh-Engel, 2008). Der wahre Wert τ einer Person zu einem Messzeitpunkt lässt sich demnach darstellen als

$$\tau = \xi + \zeta,$$

was in Kombination mit obiger Formel zu folgender Darstellung des beobachteten Wertes x führt:

$$x = \xi + \zeta + \varepsilon.$$

Dies lässt sich grafisch nach Kelava und Schermelleh-Engel (2008) wie in Abbildung 27 darstellen.

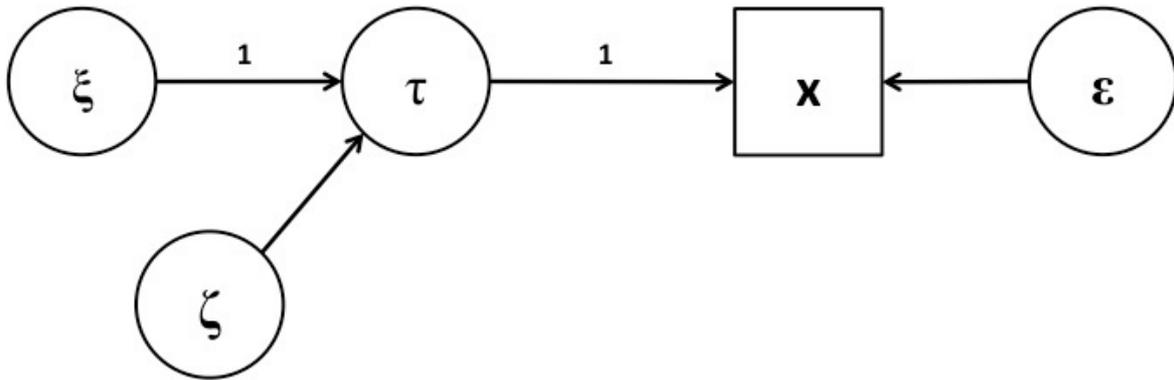


Abbildung 27: Messwert als latenter Trait-Wert und latentes State-Residuum

Der latente Trait-Wert drückt den zeitlich stabilen, also den situationsunabhängigen Anteil des beobachteten Wertes aus. Er gibt die interindividuellen Unterschiede zwischen Individuen an. Das latente State-Residuum gibt den Anteil der Messung wieder, der auf die Situation bzw. die Interaktion von Situation und Person zurückzuführen ist. In dem latenten State-Residuum drücken sich die Unterschiedlichkeit der Situation sowie das Zusammenspiel zwischen den unterschiedlichen Situationen und der Person aus.

Aus der Zerlegung des latenten State-Werts in einen latenten Trait-Wert und in ein latentes State-Residuum folgt auch die für die LST-Theorie gültige Varianzzerlegung, nach der sich die Varianz des latenten State-Werts additiv in die Varianz des latenten Trait-Werts, die Varianz des latenten State-Residuums und die Varianz des Messfehlers zerlegen lässt. Diese Varianzzerlegung ist die Grundlage zur Bestimmung verschiedener Kennwerte der LST-Theorie (Steyer et al., 1992, Steyer et al., 1999). Dies sind die *Reliabilität*, *Konsistenz* und *Spezifität* der Messwerte zu einem Zeitpunkt sowie die *Stabilität der States*, die *Stabilität der Traits* und die *Homogenität des Traits* der Personen zu zwei Zeitpunkten.

Die Reliabilität gibt den Anteil der Varianz der Variable X an, die durch interindividuelle Unterschiede, intraindividuelle Unterschiede sowie die Interaktion von Person und Situation bedingt wird. Die Konsistenz gibt den Anteil der Varianz der Variablen X an, die durch interindividuelle Unterschiede aufgrund des Traits verursacht wurden. Die Spezifität gibt den Anteil der Variablen X wieder, die durch die Unterschiede in der Situation sowie der Interaktion der Person mit der Situation bedingt wird (Steyer et al., 1999). Da die Reliabilität wie in der klassischen Testtheorie als Quotient der Varianz von wahren Werten und der Varianz der

Messwerte definiert ist, folgt zusammen mit der für die LST-Theorie gültigen Varianzzerlegung, dass sich die Reliabilität additiv aus der Konsistenz und Spezifität zusammensetzt.

Die Korrelationen der latenten State-Werte bzw. die Korrelationen der latenten Trait-Werte über zwei Messzeitpunkte geben die Stabilität der States bzw. Traits wieder. Die Korrelation der latenten Trait-Werte über zwei Messzeitpunkte und zwei Personen hinweg gibt die Homogenität des Traits wieder. Mithilfe der LST-Theorie ist es möglich, verschiedene Fragen zu beantworten (Steyer et al., 1999). Diese sind:

Wie groß ist der Anteil der Varianz, der auf den Trait (Konsistenz), auf die Situation und die Interaktion von Situation und Person (Spezifität) und den Messfehler zurückzuführen ist? Die Antwort auf diese Frage stellt den prototypischen Anwendungsfall der LST-Theorie dar. Sie gibt Auskunft darüber, inwieweit ein Messinstrument traitartige oder stateartige Eigenschaften erfasst und wie reliabel die Messung ist. Bei der Erfassung epistemologischer Überzeugungen kann die LST-Theorie zur Aufklärung der Frage beitragen, ob die zuvor postulierte stateartige Natur epistemologischer Überzeugungen sich empirisch nachweisen lässt. Insbesondere kann dies helfen, die bei Skalen zur Erfassung epistemologischer Überzeugungen häufig anzutreffenden niedrigen Reliabilitäten zu erklären. Jedoch darf der Begriff Reliabilität im Sinne der klassischen Testtheorie, wie er üblicherweise bei Skalen zur Angabe der Messgenauigkeit verwendet wird, nicht mit dem Begriff der Reliabilität aus der LST-Theorie gleichgesetzt werden. Im Rahmen der LST gliedert sich die Reliabilität in die Konsistenz und die Spezifität und gibt auf Basis dieser Zerlegung an, ob ein Konstrukt als State oder als Trait zu betrachten ist. Die Konzeption epistemologischer Überzeugungen als abhängig vom jeweiligen Kontext würde also bedeuten, dass eine hohe Spezifität und eine niedrige Konsistenz erwartet werden. Erfolgt die Messung der States und Traits reliabel, bedeutet dies, dass die Spezifität und die Konsistenz zusammen einen Großteil der Varianz des Messwertes bedingen.

Wie ist die Korrelation verschiedener Traits? Diese Frage beschäftigt sich mit den Beziehungen verschiedener Traits. Innerhalb der LST-Theorie werden States und Traits voneinander getrennt, sodass die Beziehungen zwischen mehreren Traits unabhängig von Einflüssen der Situation sowie der Interaktion von Situation und Person erfasst werden können, was bei reinen Traitmodellen nicht möglich ist. Es kann die Struktur der Zusammenhänge epistemologi-

scher Überzeugungen geprüft werden, ohne dass die Traits mit Einflüssen von States konfundiert werden.

Wie ist die Korrelation verschiedener latenter State-Residuen? Die dritte Frage beschäftigt sich mit dem Zusammenhang der latenten State-Residuen verschiedener Konstrukte. Da die latenten State-Residuen von jeglichen Einflüssen der Traits bereinigt sind, können somit deren Zusammenhänge geschätzt werden. So kann untersucht werden, wie stark der Einfluss der Situation und der Interaktion der Situation und der Person zwischen verschiedenen Konstrukten zusammenhängen.

Was sind die Korrelate von State-Änderungen? Das Thema dieser Frage sind intraindividuelle Veränderungen von States zwischen zwei Messzeitpunkten. Änderungen von States drücken situationsspezifische Effekte aus. Es besteht die Möglichkeit zu prüfen, ob die Veränderungen von exogenen Variablen oder anderen, im Modell enthaltenen Konstrukten verursacht werden. Die LST-Theorie bzw. bestimmte Modelle im Rahmen der LST-Theorie ermöglicht bzw. ermöglichen es, die Zusammenhänge der latenten State-Residuen und weiteren Variablen zu untersuchen. Dadurch kann z. B. bestimmt werden, welche Variablen systematisch mit Änderungen der latenten State-Residuen zusammenhängen. In Bezug auf epistemologische Überzeugungen könnte geprüft werden, wie hoch der Einfluss des themenspezifischen Vorwissens auf das latente State-Residuum ist. Auf diese Weise wäre eine empirische Prüfung des zuvor aufgeführten Einflusses des Vorwissens auf epistemologische Überzeugungen möglich.

Was sind die Korrelate von Trait-Änderungen? Bei dieser Frage stehen entwicklungspsychologische Betrachtungen im Mittelpunkt. Gefragt wird nach interindividuellen Unterschieden in intraindividuellen Entwicklungsverläufen zwischen zwei Messzeitpunkten. Mithilfe bestimmter LST-Modelle ist es möglich, die Variation der States von Veränderungen im Trait zu unterscheiden. Auf diese Weise können langfristige Veränderungen epistemologischer Überzeugungen getrennt von kurzfristigen Schwankungen untersucht werden.

Gelten die in den ersten fünf Fragen aufgedeckten Eigenschaften und Zusammenhänge auch für unterschiedliche Gruppen? Diese Frage beschäftigt sich mit der Geltung von LST-Modellen in unterschiedlichen Gruppen. Damit können einige der schon angesprochenen Fragestellungen untersucht werden, so unter anderem, ob es Geschlechtsunterschiede und Unterschiede zwischen verschiedenen Kulturen gibt. Diese Fragestellung erschöpft sich aber nicht

nur auf Geschlechts- und Kulturunterschiede, vielmehr können auch Unterschiede zwischen Domänen bzw. verschiedenen Fächern mithilfe dieser Methodik untersucht werden. Der Vorteil der LST-Theorie gegenüber den bereits vorgestellten Verfahren ist, dass hier explizit zwischen States und Traits unterschieden und somit die Kontextabhängigkeit epistemologischer Überzeugungen berücksichtigt wird.

Die LST-Theorie ermöglicht es also, eine Reihe von verschiedenen Fragestellungen zu untersuchen, die für die Forschung zu epistemologischen Überzeugungen relevant sind. Je nach Fragestellung wird dafür ein spezifisches LST-Modell benötigt (vgl. Steyer et al., 1999). Die ausschlaggebende Frage ist, ob sich epistemologische Überzeugungen wie von Moschner und Gruber (2005b) empirisch mit einem *Single-Trait-Multi-State-Modell* (Kelava & Schermelleh-Engel, 2008) untersuchen lassen. Im Folgenden wird dieses Modell sowie die Bedingungen, die zur Anwendung des Modells erfüllt sein müssen, vorgestellt.

Grundlegend für das Vorgehen der LST-Theorie ist, dass die benutzten Messinstrumente zu mehreren Zeitpunkten denselben Versuchspersonen vorgelegt werden. Dies ist notwendig, um die angestrebte Trennung dispositiver und situationaler Anteile der Messungen durchführen zu können. Innerhalb jedes Messzeitpunktes müssen die zu messenden Eigenschaften wiederum mit parallelen Items bzw. parallelen Tests erfasst werden. Die Anwendung paralleler Items oder Tests ist erforderlich, um die dispositionellen und situativen Einflüsse von den Messfehlern trennen zu können (Kelava & Schermelleh-Engel, 2008). Die LST-Modelle verlangen die Gültigkeit einer parallelen Äquivalenzhypothese, um die interessierenden Parameter schätzen zu können (Steyer, Ferring & Schmitt, 1992).

Mithilfe des Single-Trait-Multi-State-Modells (s. Abbildung 28) kann die Frage beantwortet werden, wie groß der Anteil der Varianz ist, der auf den Trait (Konsistenz), die Situation sowie die Interaktion von Situation und Person (Spezifität) und den Messfehler (Reliabilität) zurückzuführen ist. Diesem Modell liegt die Annahme zugrunde, dass die Traitvariable bei zwei (oder auch mehr) Messgelegenheiten und zwischen den beobachteten Variablen gleich ist. Es gibt also keine Veränderungen der Merkmalsausprägungen auf der Traitvariablen.

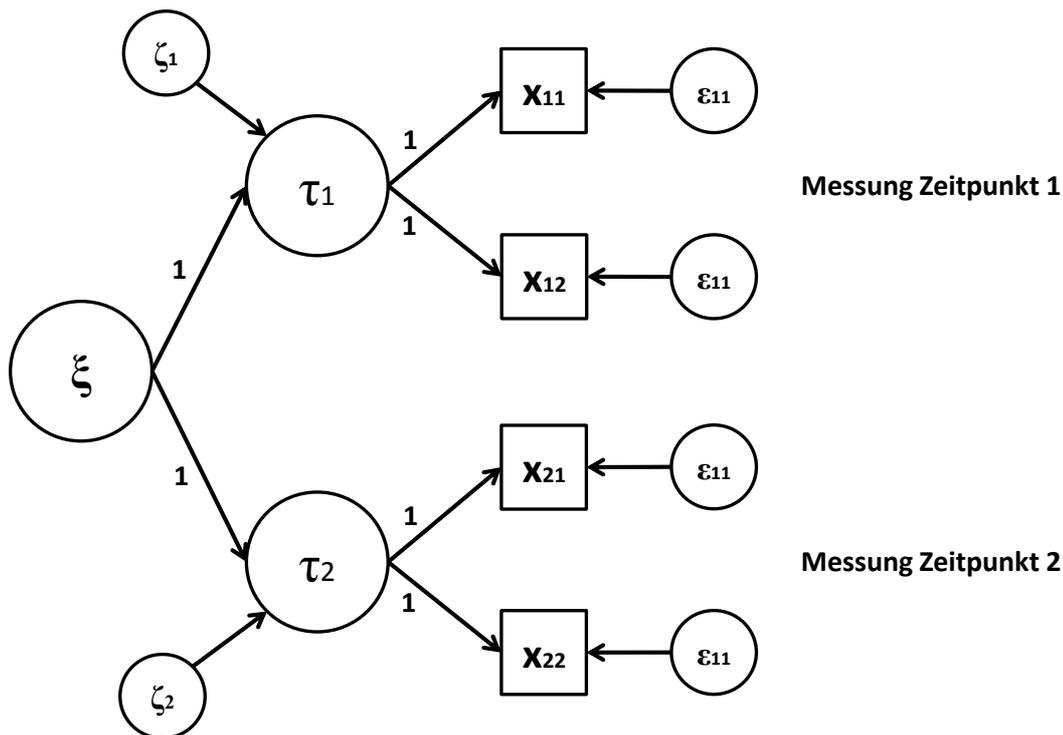


Abbildung 28: Single-Trait-Multi-State Modell

Die Konstanz der Traitvariablen wird in Abbildung 28 durch die latente Variable ξ dargestellt, die auf die beiden latenten State-Variablen τ_1 und τ_2 wirken. Die Äquivalenzannahme findet sich in den Fixierungen der Ladungen der manifesten Variablen auf den State-Variablen sowie den den Fixierungen der Ladungen der States auf der Trait-Variablen auf eins wieder. Diese Annahme muss im Fall zweier Messzeitpunkte getroffen werden, um die Identifikation des Modells zu gewährleisten (Schermelleh-Engel, Keith, Moosbrugger & Hodapp, 2004). Erweitert man das Modell auf drei Messzeitpunkte, können die Ladungen der States auf den Trait-Variablen frei geschätzt werden.

Das Single-Trait-Multi-State-Modell ermöglicht die Schätzung der Reliabilität, Konsistenz und Spezifität (Kelava & Schermelleh-Engel, 2008). Werden diese mithilfe eines Single-Trait-Multi-State-Modells geschätzt, kann die Frage beantwortet werden, ob die Latent-State-Trait-Formulierung empirisch gestützt werden kann. So kann die Passung des Modells anhand der üblichen Indizes zur Modellprüfung beurteilt werden. Wenn das Modell eine ausreichende Passung aufweist, können die Koeffizienten der Reliabilität, Konsistenz und Spezifität interpretiert werden. Allgemein ist man an einer möglichst reliablen Messung interessiert, denn nur dann erfolgt die Messung der Variablen möglichst genau. Werden epistemologische Überzeugungen als vom Kontext abhängig angesehen, würde dies bedeuten, dass bei

der Zerlegung der Reliabilität in die Komponenten Konsistenz und Spezifität ein hoher Anteil an Spezifität und ein kleiner Anteil an Konsistenz zu erwarten sind. Allerdings besteht wie bei der Berechnung der Reliabilität im Rahmen der klassischen Testtheorie keine Möglichkeit, die Höhe der Konsistenz und Spezifität zu testen, es ist also nur eine Beurteilung der numerischen Größe dieser Koeffizienten möglich.

Zusätzlich zu der Frage, ob epistemologische Überzeugungen als situationsspezifisch betrachtet werden können, können wie schon aufgeführt, Fragen zur Geschlechtsabhängigkeit und Domänenspezifität epistemologischer Überzeugungen geklärt werden. Im Rahmen von Mehrgruppenanalysen kann geprüft werden, ob sich die aufgestellten Modelle zwischen den Geschlechtern bzw. zwischen verschiedenen Domänen unterscheiden. Das grundsätzliche Vorgehen bei solchen Mehrgruppenvergleichen wurde in Kapitel 5.2 beschrieben.

Trotz der Vorteile und Reichhaltigkeit an Fragestellungen, die durch eine LST-Theorie epistemologischer Überzeugungen untersucht werden können, gibt es dennoch erhebliche methodische Probleme. Zunächst stellt sich die Problematik, dass zu den Messzeitpunkten parallele Skalen benötigt werden. Die hier durchgeführten Analysen der psychometrischen Eigenschaften der bis jetzt verfügbaren Skalen zur Messung epistemologischer Überzeugungen haben aber gezeigt, dass bei vielen Skalen diese Parallelität nicht gegeben ist. Für eine empirische Anwendung der LST-Theorie epistemologischer Überzeugungen bedarf es also der Entwicklung entsprechend paralleler Skalen. Ein weiterer problematischer Punkt innerhalb der LST-Theorie ist die Annahme, dass die faktorielle Struktur der State- und Traitvariablen jeweils gleich ist. Dies führt dazu, dass die Faktorstruktur der States für alle Individuen als gleich angenommen werden muss (Hamaker, Nesselrode & Molenaar, 2007). Die Annahme gleicher Faktorenstrukturen für alle Individuen ist aber problematisch, da von interindividuellen Strukturen nicht auf intraindividuelle Strukturen und Prozesse geschlossen werden kann und umgekehrt (Borsboom, Kievit, Cervone & Hood, 2009b). Für verschiedene psychologische Konstrukte gibt es Evidenzen, dass die Struktur der intraindividuellen Variation zwischen unterschiedlichen Individuen nicht gleich ist (Borkenau & Ostendorf, 1998, Garfein & Smyer, 1991, Hooker, Nesselrode, Nesselrode & Lerner, 1987). Die Annahmen, dass die intraindividuelle Variation bei verschiedenen Individuen gleich ist, muss selbst wiederum empirisch gezeigt werden (Hamaker et al., 2005). Empirisch wird die Annahme von gleichen Strukturen in der Form nachgewiesen, dass die Gültigkeit lokaler Homogenität (vgl. Kapitel 1.7) geprüft

wird.⁶⁰ Lässt sich die Gleichheit der faktoriellen Struktur der State- und Traitvariablen nicht nachweisen, kann z. B. auf das *Integrated-State-Trait-Modell* von Hamaker et al. (2007) zurückgegriffen werden. Dieses Modell verlangt nicht die Annahme gleicher Faktorenstrukturen und gestattet die gleichen Aussagen wie die LST-Theorie, ist allerdings in der Anwendung komplizierter als die LST-Theorie.

Ausgehend von den gestellten Forschungsfragen und den empirischen Ergebnissen konnte verdeutlicht werden, dass eine Revision der theoretischen Vorstellungen über epistemologische Überzeugungen notwendig ist. Es wurde eine Reihe von Befunden vorgestellt, die in der Literatur über epistemologische Überzeugungen häufig festzustellen sind. Dies sind vor allem die schlechte Replizierbarkeit der Faktorenstrukturen von Items zur Erfassung epistemologischer Überzeugungen sowie die niedrigen Reliabilitäten entsprechender Skalen. Da diese Befunde im Zusammenhang mit der persönlichkeitspsychologischen Sichtweise epistemologischer Überzeugungen stehen, stellt sich die Frage, ob diese nicht aufgegeben werden sollte. Hier wird die Position vertreten, dass die persönlichkeitspsychologische Sichtweise beibehalten werden soll, wobei allerdings ein Perspektivenwechsel innerhalb dieses Ansatzes erforderlich ist. Die sich hieraus ergebenden Perspektiven sollen zum Abschluss kurz angerissen werden.

Der persönlichkeitspsychologische Ansatz, der im Rahmen der Forschung über epistemologische Überzeugungen bis jetzt hauptsächlich herangezogen wurde, hat seinen Schwerpunkt hauptsächlich auf der Analyse interindividueller Differenzen. Wie aber angeführt wurde, können durch Untersuchung von interindividuellen Differenzen keine Rückschlüsse auf intraindividuelle Unterschiede gezogen werden. Es gibt keinen allgemeinen eindeutigen Zusammenhang zwischen inter- und intraindividuellen Differenzen (Borsboom et al., 2009b, Molenaar, Huizenga & Nesselroade, 2003). Da die intraindividuelle Variation, welche die Situationsabhängigkeit als konstatierendes Element epistemologischer Überzeugungen herausstellt, gemäß der zu dargelegten Argumentation essenziell für deren Konzeption ist, kann eine einfache Betrachtung interindividueller Differenzen nicht ausreichend sein. Nach Cervone (2005) lie-

⁶⁰ Neben der Annahme lokaler Homogenität ist zusätzlich die Geltung von Ergodizität nachzuweisen, erst dann kann von einer vollständigen Äquivalenz der inter- und intraindividuellen Strukturen ausgegangen werden (Borsboom, Kievit, Cervone & Hood, 2009). Ergodizität besagt, dass die Ergebnisse der Untersuchung von n Personen, wenn n gegen unendlich geht, gleich den Ergebnissen der Untersuchung einer Zeitreihe von t Zeitpunkten, wenn t gegen unendlich geht, sein müssen. Der Nachweis von Ergodizität verlangt zwei Teilbedingungen: Zum einen müssen die Beobachtungen stationär sein, d. h. die Beobachtungen müssen stabile statistische Kennwerte, wie z. B. einen konstanten Mittelwert, aufweisen und zum anderen müssen alle Beobachtungen gleiche Trajektorien aufweisen (Molenaar, 2004).

fert die Betrachtung interindividueller Differenzen nur eine Taxonomie der Struktur individueller Unterschiede, die aber für eine Theorie epistemologischer Überzeugungen wenig nützlich ist. Diese Taxonomie ist lediglich eine Sichtweise, die zur Klassifikation von Personen genutzt werden kann. Somit handelt es sich bei den Konstrukten, die interindividuelle Unterschiede beschreiben, um klassifizierende Konstrukte. Diese sind aber nach Cervone (2005) vom falschen logischen Typ, um als Erklärung für das Erleben und Verhalten einer Person dienen zu können. Die andere – in der Persönlichkeitspsychologie weniger beachtete – Sichtweise ist die Betrachtung intraindividuelle Strukturen und Prozesse. Hier wird also nach einer Persönlichkeitsarchitektur gefragt, welche die Zusammenhängestrukturen und Prozesse von Persönlichkeitsvariablen in einer intraindividuellen Perspektive betrachtet (Cervone, 2004). Damit einhergehend ist auch ein methodischer Perspektivenwechsel verbunden, der von der querschnittlichen Untersuchung einer Menge von Personen absieht und auf die längsschnittliche Untersuchung einer oder mehrerer Personen abzielt, um intraindividuelle Strukturen und Prozesse aufdecken zu können (Molenaar, 2004, Molenaar & Campbell, 2009). Der persönlichkeits-theoretische Ansatz epistemologischer Überzeugungen ist daher nicht aufzugeben, sondern durch einen Ansatz zu ergänzen, der die Dynamik epistemologischer Überzeugungen berücksichtigt.

7. Literaturverzeichnis

- Ajzen, I. & Fishbein, M. (1977). Attitude-behavior relations: a theoretical analysis and review of empirical research. *Psychological Bulletin*, 84, 888-918.
- Allison, P. (2009). Missing Data. In R. E. Millsap & A. Maydeu-Olivares (Hrsg.), *The SAGE Handbook of Quantitative Methods in Psychology* (S. 72-89). Thousand Oaks: Sage
- Allison, P. (2012, April). *Handling Missing Data by Maximum Likelihood*. Keynote presentation at the SAS Global Forum, Orlando, Florida
- Amelang, M. & Zienlinski, W. (2002). *Psychologische Diagnostik und Intervention*. Heidelberg: Springer.
- Artelt, C. (2000). Wie prädiktiv sind retrospektive Selbstberichte über den Gebrauch von Lernstrategien für strategisches Lernen? *Zeitschrift für Pädagogische Psychologie*, 14, 72–84.
- Asendorpf, J. (1995). Persönlichkeitspsychologie: Das empirische Studium der individuellen Besonderheit aus spezieller und differentieller Perspektive. *Psychologische Rundschau*, 46, 235-247.
- Asendorpf, J. (1999). *Persönlichkeitspsychologie*. Heidelberg: Springer.
- Asparouhov, T. & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, 16, 397-438.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2003). *Multivariate Analysemethoden*. Heidelberg: Springer.
- Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling*, 9, 78-102.
- Bandalos, D. L. & Finney, S. J. (2001). Item parceling issues in structural equation modeling. In G. A. Marcoulides & R. E. Schumacker (Hrsg.), *New developments and techniques in structural equation modeling* (S. 269-296). Mahwah: Lawrence Erlbaum.
- Barca, A. (1999). *Manual del cuestionario des procesos y estrategias des aprendizaje para el alumnado de educación secundaria*. La Coruña: Publicaciones de la Revista Galego-Portuguesa des Psícoloxía e Educación.
- Baron, J. (1993). Why teach thinking? – An essay. *Applied Psychology: An International Review*, 42, 191-237.
- Bartlett, F. C. (1932). *Remembering: An experimental and social study*. Cambridge: Cambridge University Press.

- Bauersfeld, H. (1988). Interaction, construction and knowledge: alternative perspectives for mathematics education. In D. A. Grouws, T. J. Cooney & D. Jones (Hrsg.), *Effective mathematics teaching* (S. 27-46). Reston: NCTM & Lawrence Erlbaum.
- Baxter Magolda, M. B. (1992). *Knowing and reasoning in college: Gender-related patterns in students' intellectual development*. San Francisco: Jossey Bass.
- Beauducel, A. (2001). On the generalizability of factors: The influence of changing contexts of variables on different methods of factor extraction. *Methods of Psychological Research Online*, 6, 69-96.
- Belenky, M. F., Clinchy, B. M., Goldberger, N. R. & Tarule, J. M. (1986). *Women's ways of knowing: The development of self, voice and mind*. New York: Basic Books.
- Bell, P. & Linn, M. C. (2002). Beliefs about science: How does science instruction contribute. In B. K. Hofer & P. R. Pintrich (Hrsg.), *Personal Epistemology. The psychology of beliefs about knowledge and knowing* (S. 321-346). Mahwah: Lawrence Erlbaum.
- Bendixen, L. B., Dunkle, M. E. & Schraw, G. (1994). Epistemological beliefs and reflective judgement. *Psychological Review*, 75, 1595-1600.
- Bendixen, L. D & Feucht, F. C. (2010). *Personal Epistemology in the Classroom: Theory, Research, and Implications for Practice*. Cambridge: Cambridge University Press.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
- Bentler, P. M. (1995). *EQS structural equations program manual*. Encino: Multivariate Software.
- Bentler, P. M. & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Biggs, J. (1987). *The study process questionnaire (SPQ): Manual*. Hawthorne: Australian Council for Educational Research.
- Biggs, J. (1993). What do inventories of students' learning really measure? A theoretical review and clarification. *British Journal of Educational Psychology*, 63, 3-19.
- Biggs, J., Kember, D. & Leung, D. Y. P. (2001). The revised two factor Study Process Questionnaire: R-SPQ-2F. *British Journal of Educational Psychology*, 71, 133-149.
- Biglan, A. (1973a). The characteristics of subject matter in different academic areas. *Journal of Applied Psychology*, 57, 195-203.
- Biglan, A. (1973b). Relationships between subject matter characteristics and the structure and output of university departments. *Journal of Applied Psychology*, 57, 204-213.
- Bischof, N. (2008). *Psychologie. Ein Grundkurs für Anspruchsvolle*. Stuttgart: Kohlhammer.

- Bless, N. & Schwarz, N. (2002). Konzeptgesteuerte Informationsverarbeitung. In D. Frey & M. Irle (Hrsg.), *Theorien der Sozialpsychologie. Band III: Motivations-, Selbst- und Informationsverarbeitungstheorien* (S. 257-278). Bern: Verlag Hans Huber.
- Boerner, S., Seeber, G., Keller, H. & Beinborn, P. (2005). Lernstrategien und Lernerfolg im Studium: Zur Validierung des LIST bei berufstätigen Studierenden. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 37, 17-26.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bollen, K. A. & Stine, R. A. (1993). Bootstrapping goodness-of-fit measures in structural equation modelling. In K. A. Bollen & J. S. Long (Hrsg.), *Testing structural equation models* (S. 111-135). Thousand Oaks: Sage.
- Borg, I. (1992). *Grundlagen und Ergebnisse der Facettentheorie*. Bern: Huber.
- Borkenau, P. & Ostendorf, F. (1998). The Big Five as states: How useful is the Five Factor model to describe intraindividual variations over time? *Journal of Research in Personality*, 32, 202–221.
- Bornstein, R. F. (2011). Toward a process-focused model of test score validity: Improving psychological assessment in science and practice. *Psychological Assessment*, 23, 532-544.
- Borsboom, D. (2008). Latent variable theory. *Measurement*, 6, 25-53.
- Borsboom, D., Cramer, A. O. J., Kievit, R. A., Zand Scholten, A. & Franic, S. (2009a). *The end of construct validity*. In R. W. Lissit (Hrsg.), *The concept of validity: Revisions, new directions, and applications* (S. 135-170). Charlotte: Information Age Publishers.
- Borsboom, D., Kievit, R. A., Cervone, D. P. & Hood, S. B. (2009b). The two disciplines of scientific psychology, or: The disunity of psychology as a working hypothesis. In J. Valsiner, P. C. M. Molenaar, M. C. D. P. Lyra & N. Chaudary (Hrsg.), *Developmental process methodology in the social and developmental sciences* (S. 67-97). New York: Springer.
- Borsboom, D., Mellenbergh, G. J. & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, 110, 203-219.
- Borsboom, D., Mellenbergh, G. J. & van Heerden, J. (2004). The concept of validity. *Psychological Review*, 111, 1061– 1071.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler*. Heidelberg: Springer.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Heidelberg: Springer.

- Brayant, F. B. (2000). Assessing the validity of measurement. In L. G. Grimm & P. R. Yarnold (Hrsg.), *Reading and understanding more multivariate statistics* (S. 99-146). Washington: APA Press.
- Brayant, F. B. & Yarnold, P. R. (2001). Principal component analysis and exploratory and confirmatory factor analysis. In L. G. Grimm & P. R. Yarnold (Hrsg.), *Reading and understanding multivariate statistics* (S. 99-136). Washington: APA Press.
- Briggs, S. R. & Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality*, 54, 106-148.
- Bromme, R. (2005). Thinking and knowing about knowledge: A plea for and critical remarks on psychological research programs on epistemological beliefs. In M. Hoffmann, J. Lenhard & F. Seeger (Hrsg.), *Activity and sign - Grounding mathematics education* (S. 191-201). New York: Springer.
- Bromme, R., Kienhues, D. & Porsch, T. (2010). Who knows what and who can we believe? Epistemological beliefs are beliefs about knowledge (mostly) attained from others. In L. D. Bendixen & F. C. Feucht (Hrsg.), *Personal Epistemology in the Classroom: Theory, Research, and Implications for Practice* (S. 163-193). Cambridge: Cambridge University Press.
- Bromme, R., Kienhues, D. & Stahl, E. (2008). Knowledge and epistemological beliefs: An intimate but complicate relationship. In M. S. Khine (Hrsg.), *Knowing, Knowledge, and Beliefs: Epistemological Studies Across Diverse Cultures* (S. 423-441). New York: Springer.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Hrsg.), *Testing structural equation models* (S. 136-162). Newbury Park: Sage.
- Buehl, M. M. (2008). Assessing the multidimensionality of students' epistemological beliefs. In M. S. Khine (Hrsg.), *Knowing, Knowledge, and Beliefs: Epistemological Studies Across Diverse Cultures* (S. 65-112). New York: Springer.
- Buehl, M. M. & Alexander, P.A. (2006). Examining the dual nature of epistemological beliefs. *International Journal of Educational Research*, 45, 28-42.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson.
- Bunge, M. (1987). *Kausalität: Geschichte und Probleme*. Tübingen: Mohr.

- Burr, J. E. & Hofer, B. K. (2002). Personal epistemology and theory of mind: Deciphering young children's beliefs about knowledge and knowing. *New Ideas in Psychology*, 20, 199–224.
- Byrne, B. M. (1998). *Structural equation modelling with LISREL, PRELIS, and SIMPLIS: Basic concepts, application, and programming*. Mahwah: Lawrence Erlbaum.
- Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, 54, 297-312.
- Cano, F. (2005). Epistemological beliefs, approaches to learning, and academic performance. *British Journal of Educational Psychology*, 75, 203-221.
- Caramazza, V. (2010). *Unveröffentlichte Magisterarbeit im Fachbereichs Erziehungswissenschaft*. Saarbrücken: Universität des Saarlandes.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1, 245-276.
- Cervone, D. (2004). The architecture of personality. *Psychological Review*, 111, 183-204.
- Cervone, D. (2005). Personality architecture: Within-person structures and processes. *Annual Review of Psychology*, 56, 423-452.
- Chapman, M. (1988). *Constructive Evolution: Origins and Development of Piaget's Thought*. Cambridge: Cambridge University Press.
- Clarebout, G., Elen, J., Luyten, L. & Bamps, H. (2001). Assessing epistemological beliefs: Schommer's questionnaire revisited. *Educational Research and Evaluation*, 7, 53-77.
- Chinn, C., Buckland, L. & Samarapungavan, A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology, *Educational Psychologist*, 46, 141-167.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum.
- Cole, R. P., Goetz, E. T. & Willson, V. (2000). Epistemological beliefs of underprepared college students. *Journal of College Reading and Learning*, 31, 60-72.
- Conley, A. M., Pintrich, P. R., Vekiri, I. & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology*, 29, 186-204.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, 78, 98-104.
- Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition*, 31, 187-276.

- Cosmides, L. & Tooby, J. (1992). Cognitive adaptations for social exchange. In J. Barkow, L. Cosmides, & J. Tooby (Hrsg.). *The adapted mind* (S. 163-228). New York: Oxford University Press.
- Costello, A. B. & Osborne, J. W. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, 10. Abgerufen am 22.03.2009. Verfügbar unter <http://pareonline.net/getvn.asp?v=10&n=7>.
- Courvoisier, D. S., Eid, M. & Nussbeck, F. W. (2007). Mixture Distribution State-Trait Models. Basic ideas and applications. *Psychological Methods*, 12, 80-104.
- Cronbach, J. L. & Mehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.
- Crystal, D. (1993). *Die Cambridge Enzyklopädie der Sprache*. Frankfurt: Campus.
- Dahl, T. I., Bals, M. & Turi, A. L. (2005). Are students' beliefs about knowledge learning associated with their reported use of learning strategies? *British Journal of Educational Psychology*, 75, 257-273.
- Danovitch, J. H. & Keil, F. C. (2004). Should I ask a fisherman or a biologist? Developmental shifts in ways of clustering knowledge. *Child Development*, 75, 918-931.
- De Jong, T. & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31, 105–113.
- DeBacker, T. K., Crowson, H. M., Beesley, A. D., Thoma, S. J. & Hestevold, N. L. (2008). The challenge of measuring epistemic beliefs: an analysis of three self-report instruments. *The Journal of Experimental Education*, 76, 281–312.
- DeRaad, B. (2000). *The big five personality factors: The psycholexical approach to personality*. Göttingen: Hogrefe und Huber.
- DeShon, R. P. (1998). A cautionary note on measurement error corrections in structural equation models. *Psychological Methods*, 3, 412-423.
- Devitt, M. (1991). *Realism and truth*. Cambridge: Blackwell.
- Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, 73, 1246–1256.
- diSessa, A. (1993). J's Epistemological Stance and Strategies. *Paper presented at the American Education Research Association*.
- diSessa, A. A., Elby, A. & Hammer, D. (2002). J's Epistemological Stance and Strategies. In G. M. Sinatra & P. R. Pintrich (Hrsg.), *Intentional conceptual change* (S. 237-290). Mahwah: Lawrence Erlbaum.

- DiStefano, C. & Motl, R. W. (2009). Personality correlates of method effects due to negatively worded items on the Rosenberg Self-Esteem scale. *Personality and Individual Differences*, 46, 309-313.
- Donald, J. G. (1990). University professors' views of knowledge and validation processes. *Journal of Educational Psychology*, 82, 242-249.
- Dweck, C. S. & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256-273.
- Eckey, H.-F., Kosfeld, R. & Rengers, M. (2002). *Multivariate Statistik*. Wiesbaden: Gabler.
- Edwards, J. R. & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures, *Psychological Methods*, 5, 2155-2174.
- Eid, M. (1997). Happiness and satisfaction: An application of a latent state-trait model for ordinal variables. In J. Rost & R. Langeheine (Hrsg.), *Applications of latent trait and latent class models in the social sciences* (S. 145-151). Münster: Waxmann.
- Elby, A. & Hammer, D. (2001). On the substance of a sophisticated epistemology. *Science Education*, 85, 554-567.
- Elder, A. D. (2002). Characterizing fifth grade students' epistemological beliefs in science. In B. K. Hofer, B. K. & Pintrich, P. R. (Hrsg.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (S. 103-118). Mahwah: Lawrence Erlbaum.
- Ellis, J. L. & Van den Wollenberg, A. L. (1993). Local homogeneity in latent trait models: A characterization of the homogeneous monotone IRT model. *Psychometrika*, 58, 417-429.
- Embretson, S. E. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin*, 93, 179-197.
- Embretson, S. E. (1994). Applications of cognitive design systems to test development. In C.R. Reynolds (Hrsg.), *Cognitive assessment: A multidisciplinary perspective*. (S. 107-135). New York: Plenum Press.
- Embretson, S. E. (1998). A cognitive design system approach for generating valid tests: Approaches to abstract reasoning. *Psychological Methods*, 3, 300-396.
- Ericsson, K. A. & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*: Cambridge: MIT Press.
- Estes, D., Chandler, M., Horvath, K. J. & Backus, D.W. (2003). American and British college students' epistemological beliefs about research on psychological and biological development. *Applied Developmental Psychology*, 23, 625-642.

- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272-299.
- Ferguson, L.E., Bråten, I., & Strømsø, H.I. (2012). Epistemic cognition and change when students read multiple documents containing conflicting scientific evidence: A think-aloud study. *Learning and Instruction*, 22, 103-120.
- Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87, 477–531.
- Fiske, S.T. & Taylor, S. E. (1991). *Social cognition*. New York: McGraw-Hill
- Fisseni, H.-J. (2004) *Lehrbuch der Psychologischen Diagnostik: Mit Hinweisen zur Intervention*. Göttingen: Hogrefe.
- Flora, D.B., & Curran, P.J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9, 466-491.
- Frederiksen, N. (1984). Implications of cognitive theory for instruction in problem solving. *Review of Educational Research*, 54, 636–407.
- Gardenne, V. (2004). *Philosophie der Psychologie*. Bern: Hans Huber.
- Garfein, A. J. & Smyer, M. A. (1991). P-technique factor analysis of the multiple affect adjective check list (MAACL). *Journal of Psychopathology and Behavioral Assessment*, 13, 155–171.
- Gay, G. (1978). Viewing the pluralistic classroom as a cultural microcosm. *Educational Research Quarterly*, 2, 45–49.
- Gerber, J. (2004). *Intergenerationale Transmission epistemologischer Überzeugungen. Entwicklung eines Erhebungsinstruments und Befunde zur Genese wissensbezogener Vorstellungen*. Dissertation, Universität Bielefeld. Abgerufen am 30.09.2007. Verfügbar unter <http://deposit.ddb.de/cgi-bin/dokserv?idn=979431670>.
- Gerbing, D. W. & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, 25, 186-192.
- Gignac, G. E. (2007). Multi-factor modeling in individual differences research: Some recommendations and suggestions. *Personality and Individual Differences*, 42, 37–48.
- Graham, J. M. (2006). Congeneric and (essentially) tau-equivalent estimates of core reliability. What they are and how to use them. *Educational and Psychological Measurement*, 66, 930-944.

- Graumann, C. F. (2000). Kontext als Problem der Psychologie. *Zeitschrift für Psychologie*, 208, 55-71.
- Green, S.A. & Yang, Y. (2009). Commentary on coefficient alpha: a cautionary tale. *Psychometrika*, 74, 121-135.
- Greene, J. A., Azevedo, R. & Torney-Purta, J. (2008). Modeling epistemic and ontological cognition: Philosophical perspectives and methodological directions. *Educational Psychologist*, 43, 142-160.
- Gruber, H. & Moschner, B. (2003). Begabung – Lernen – Wissen: Epistemologische Überzeugungen. *Unpubliziertes Manuskript, Universität Regensburg*.
- Gruber, H. & Stamouli, E. (2009). Intelligenz und Vorwissen. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 28-47). Heidelberg: Springer.
- Gubler, H. & Bischof, N. (1993). Computerspiele als Werkzeug der motivationspsychologischen Grundlagenforschung. *Zeitschrift für Psychologie*, 201, 287-315. Hagemann, D., Hewig, J., Seifert, J., Naumann, E. & Bartussek, D. (2005). The latent state-trait structure of resting EEG asymmetry: Replication and extension. *Psychophysiology*, 42, 740-752.
- Hall, R. J., Snell, A. F. & Foust, M. S. (1999). Item parceling strategies in SEM: Investigating the subtle effects of unmodeled secondary constructs. *Organizational Research Methods*, 2, 233-256.
- Hamaker, E. L., Dolan, C. V. & Molenaar, P. C. M. (2005). Statistical modeling of the individual: Rationale and application of multivariate time series analysis. *Multivariate Behavioral Research*, 40, 207–233.
- Hamaker, E. L., Nesselroade, J. R. & Molenaar, P. C. M. (2007). The integrated trait-state model. *Journal of Research in Personality*, 41, 295-315.
- Hammer, D. (1989). Two approaches to learning physics. *The Physics Teacher*, 27, 664-670.
- Hammer, D. & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer & P.R. Pintrich (Hrsg.), *Personal Epistemology. The psychology of beliefs about knowledge and knowing* (S. 169-190). Mahwah: Lawrence Earlbaum.
- Hartig, J., Frey, A. & Jude, N. (2008). Validität. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 136-163). Heidelberg: Springer.
- Hartung, J. & Elpelt, B. (1999). *Multivariate Statistik*. München: Oldenbourg.
- Hasselhorn, M. & Gold, A. (2006). *Pädagogische Psychologie*. Stuttgart: Kohlhammer.
- Heckhausen, H. (1989). *Motivation und Handeln*. Berlin: Springer.
- Herzog, W. (2012). *Wissenschaftstheoretische Grundlagen der Psychologie*. Wiesbaden: VS Verlag für Sozialwissenschaften.

- Hildebrandt, L. & Temme, D. (2006). Probleme der Validierung mit Strukturgleichungsmodelle. *SFB 649 Discussion Paper*. Abgerufen am 08.09.2008. Verfügbar unter <http://sfb649.wiwi.hu-berlin.de>.
- Hodapp, V. (1994). *Analyse linearer Kausalmodelle*. Bern: Hans Huber.
- Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology*, 25, 378-405.
- Hofer, B. K. (2001). Personal epistemology research: Implications for learning and transfer. *Educational Psychology Review*, 13, 353-383.
- Hofer, B. K. (2002). Personal Epistemology as a psychological and educational construct: An introduction. In B. K. Hofer & P.R. Pintrich (Hrsg.), *Personal Epistemology. The psychology of beliefs about knowledge and knowing* (S. 3-14). Mahwah: Lawrence Erlbaum.
- Hofer, B. K. (2004a). Introduction: Paradigmatic Approaches to Personal Epistemology. *Educational Psychologist*, 39, 1-3.
- Hofer, B. K. (2004b). Epistemological understanding as a metacognitive process: Thinking aloud during online searching. *Educational Psychologist*, 39, 43-55.
- Hofer, B. K. & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67, 88-140.
- Hofer, B. K. & Pintrich, P. R. (2002). *Personal epistemology. The psychology of beliefs about knowledge and knowing*. Mahwah: Lawrence Erlbaum.
- Hofer, B. K. & Sinatra, G. M. (2010). Epistemology, metacognition, and self-regulation: musings on an emerging field. *Metacognition and Learning*, 5, 113-120.
- Hooker, K., Nesselroade, D. W., Nesselroade, J. R. & Lerner, R. M. (1987). The structure of intraindividual temperament in the context of mother-child dyads: P-technique factor analyses of short-term change. *Developmental Psychology*, 23, 332-346.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179-185.
- Hu, L., & Bentler, P. (1995). Evaluating model fit. In R. H. Hoyle (Hrsg.), *Structural equation modeling. Concepts, issues, and applications* (S. 76-99). London: Sage.
- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge: MIT Press.
- Jansen, B. R. J. & Van der Maas, H. L. J. (1997). Statistical tests of the rule assessment methodology by latent class analysis. *Developmental Review*, 17, 321-357.

- Jansen, B. R. J. & Van der Maas, H. L. J. (2002). The development of children's rule use on the balance scale task. *Journal of Experimental Child Psychology*, 81, 383–416.
- Jehng, J. J. (1991). *The nature of epistemological beliefs about learning*. Unveröffentlichte Dissertation, Universität Illinois, Urbana-Champaign.
- Jehng, J. J., Johnson, S. D. & Anderson, R. C. (1993). Schooling and students' epistemological beliefs about learning. *Contemporary Educational Psychology*, 18, 23–35.
- Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, 36, 109-133.
- Jöreskog, K. G. (1993). Testing structural equation models. In K. A. Bollen & J. S. Long (Hrsg.), *Testing structural equation models* (S. 294-317). Newbury Park: Sage.
- Jöreskog, K. G., & Sörbom, D. (1993). *Structural equation modeling with the SIMPLIS command language*. Chicago: Scientific Software.
- Kaiser, H. F. & Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*, 34, 11-117.
- Keil, F. C., Stein, C., Webb, L., Billings, V. D. & Rozenblit, L. (2008). Discerning the division of cognitive labor: An emerging understanding of how knowledge is clustered in other minds. *Cognitive Science*, 32, 259–300.
- Kelava, A. & Schermelleh-Engel, K. (2008). Latent-State-Trait-Theorie (LST-Theorie). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 344-360). Heidelberg: Springer.
- Kember, D., Leung, D., Jones, A., Loke, A. J., MacKay, J., Sinclair, K., Tse, H., Webb, C., Wong, F. K. Y., Wong M. & Young, E. (2000). Development of a questionnaire to measure the level of reflective thinking. *Assessment and Evaluation in Higher Education*, 25, 381-389.
- Kennedy, J. C. (2007, August). The effect on model fit of ignoring item wording effects in the core self-evaluations scale. *Paper Accepted for 2007 Annual Meeting of the Academy of Management, Philadelphia*.
- Khine, M. S. (2008) *Knowing, knowledge, and beliefs: Epistemological studies across diverse cultures*. New York: Springer.
- Kienhues, D., Bromme, R. & Stahl, E. (2008). Changing epistemological beliefs: The unexpected impact of a short-term intervention. *British Journal of Educational Psychology*, 78, 545-565.

- King, P. M., Kitchener, K. S., Davison, M. L., Parker, C. A., & Wood, P. K. (1983). The justification of beliefs in young adults: A longitudinal study. *Human Development*, 26, 106-116.
- King, P. M., Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco: Jossey Bass.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95, 163-182.
- Kintsch, W. (1998). *Comprehension: A Paradigm for Cognition*. Cambridge: Cambridge University Press.
- Kitchener, K. S. (1983). Cognition, metacognition, and epistemic cognition. *Human development*, 26, 222-232.
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford Press.
- Klopp, E. (2010). Explorative Faktorenanalyse. Verfügbar unter: <http://psydok.sulb.uni-saarland.de/volltexte/2013/4823/>. Zuletzt abgerufen am 07.08.2013.
- Klopp, E. & Stark, R. (in Druck). Quantitative und qualitative Analysen zur Validität und Reliabilität des Fragebogens zur Erfassung epistemologischer Überzeugungen (FEE). In A. Anschütz, H. Gruber & B. Moschner (Hrsg.), *Wissen und Lernen - in der Sicht von Lernenden und Lehrenden*. Münster: Waxmann.
- Knorr Cetina, K. D. (1999). *Epistemic Cultures. How the Sciences Make Knowledge*. Cambridge: Harvard University Press.
- Köller, O., Baumert, J. & Neubrand, J. (2000). Epistemologische Überzeugungen und Fachverständnis im Mathematik- und Physikunterricht. In J. Baumert, W. Bos & R. Lehmann (Hrsg.), *TIMSS/III: Dritte Internationale Mathematik- und Naturwissenschaftsstudie - Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn*. Opladen: Leske+Budrich.
- Kopp, B. & Mandl, H. (2005). *Wissensschemata*. In H. Mandl & H. F. Friedrich (Hrsg.), *Handbuch Lernstrategien* (S. 38-49). Göttingen: Hogrefe.
- Krapp, A. (2001). Interesse. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie*. Weinheim: BeltzPVU, 286-294.
- Krause, U.-M. & Stark, R. (2006). Vorwissen aktivieren. In H. Mandl & H. F. Friedrich (Hrsg.), *Handbuch Lernstrategien* (S. 38-49). Göttingen: Hogrefe.

- Krettenauer, T. (2005). Die Erfassung des Entwicklungsniveaus epistemologischer Überzeugungen und das Problem der Übertragbarkeit von Interviewverfahren in standardisierte Fragebogenmethoden. *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie*, 37, 69-79.
- Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, 50, 537-567.
- Kuhn, D. (1999). Metacognitive development. In L. Balter & C. Tamis Le-Monda (Hrsg.), *Child psychology: A handbook of contemporary issues*. Philadelphia: Psychology Press.
- Kuhn, D. (1991). *The Skills of Argument*. Cambridge: Cambridge University Press.
- Kyngdon, A. (2013). Descriptive theories of behavior may allow for the psychological measurement of attributes. *Theory and Psychology*, 23, 227-250.
- Leach, J., Millar, R., Ryder, J. & Séré, M.-G. (2000). Epistemological understanding in science learning: The consistency of representations across contexts. *Learning and Instruction*, 10, 497-527.
- Lee, J. J. (2012). Correlation and causation in the study of personality. *European Journal of Personality*, 26, 372-390.
- Leung, K. & Bond, M. H. (1989). On the empirical identification of dimensions for cross-cultural comparisons. *Journal of Cross Cultural Psychology*, 20, 133-151.
- Limón, M. (2006a). Introduction. *International Journal of Educational Research*, 45, 1-6.
- Limón, M. (2006b). The domain generality-specificity of epistemological beliefs: A theoretical problem, a methodological problem or both? *International Journal of Educational Research*, 45, 7-27.
- Little, T. D., Cunningham, W. A., Shahar, G. & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, 9, 151-73.
- Loehlin, J. C. (1987). *Latent variable models*. Hillsdale: Lawrence Earlbaum.
- Loftus, E. F. (1975). Spreading activation within semantic categories: Comments on Rosch's „Cognitive representation of semantic categories.“ *Journal of Experimental Psychology: General*, 104, 234-240.
- Lonka, K. & Linblom-Yläne, S. (1996). Epistemologies, conceptions of learning, and study practices in medicine and psychology. *Higher Education*, 31, 5-24.
- Lord, F. M. & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading: Addison-Wesley.

- Louca, L., Elby, A., Hammer, D. & Kagey, T. (2004) Epistemological resources: Applying a new epistemological framework to science instruction. *Educational Psychologist*, 39, 57-68.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. *Psychologische Rundschau*, 58, 103-117.
- Lutz, D. J. & Keil, F. C. (2002). Early understanding of the division of cognitive labor. *Child Development*, 73, 1073-1084.
- MacCallum, R. C. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin*, 100, 107–120.
- MacCallum, R. C., Roznowski, M. & Necowitz, L.B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, 111, 490–504.
- MacCallum, R. C., Wegener, D. T., Uchino, B., N. & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin*, 114, 185–199.
- MacCallum, R. C., Widaman, K. F., Zhang, S. & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4, 87-99.
- Maggioni, L., Alexander, P. & van Sledright, B. (2004). At a crossroads: the development of epistemological beliefs in historical thinking. *European Journal of School Psychology*, 2, 169-200.
- Mardia, K.V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57, 519–530.
- Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifactors? *Journal of Personality and Social Psychology*, 70, 810-819.
- Marsh, H. W. (2007). Application of confirmatory factor analysis and structural equation modeling in sport and exercise psychology. In G. Tenenbaum & R. C. Eklund (Hrsg.), *Handbook of Sport Psychology* (S. 744-798). New Jersey: Wiley & Sons.
- Marsh, H. W., Hau, K. T., Balla, J. R. & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research*, 33, 181-220.
- Marton, F. & Säljö, R. (1976). On qualitative differences in learning – I: Outcome and process. *British Journal of Educational Psychology*, 46, 4-11.

- Mason, L. & Boldrin, A. (2008). Epistemic metacognition in the context of information searching on the web. In M. S. (Hrsg.), *Knowing, Knowledge, and Beliefs: Epistemological Studies Across Diverse Cultures* (S. 377-404). New York: Springer.
- Mason, L. & Bromme, R. (2010). Situating and relating epistemological beliefs into metacognition: studies on beliefs about knowledge and knowing. *Metacognition and Learning*, 5, 1-6.
- Mayring, P. (1996). *Einführung in die qualitative Sozialforschung. Eine Anleitung zu qualitativem Denken*. Weinheim: PVU.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah: Earlbaum.
- McDonald, R. P. & Ho, M.-H. R. (2002). Principles and practice in reporting structural equation analysis. *Psychological Methods*, 7, 64-82.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58, 525-543.
- Messick, S. (1989). Validity. In R. L. Linn (Hrsg.), *Educational measurement* (S. 13-103). Washington, DC: American Council on Education and National Council on Measurement in Education.
- Meyer, H. (2004). *Theorie und Qualitätsbeurteilung psychometrischer Tests*. Stuttgart: Kohlhammer.
- Mezirow, J. (1991). *Transformative dimensions of adult learning*. San Francisco: Jossey-Bass.
- Mezirow, J. (1998). On critical reflection. *Adult Education Quarterly*, 48, 185-198.
- Minsky, M. (1986). *The society of mind*. New York: Simon & Schuster.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- Mislevy, R. (2008). How cognitive science challenges the educational measurement tradition. *Measurement*, 6, 124.
- Molenaar, P. C. M. (2004). A manifesto on Psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, 2, 201-218.
- Molenaar, P. C. M. & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychology*, 18, 112-117.
- Molenaar, P. C. M., Huizenga, H. M. & Nesselroade, J. R. (2003). The relationship between the structure of interindividual and intraindividual variability: A theoretical and empirical vindication of developmental systems theory. In U. M. Staudinger & U. Lindenberger (Hrsg.), *Understanding human development: Dialogues with lifespan psychology* (S. 339-360). Dordrecht: Kluwer Academic Publishers.

- Montada, L. (2008). Fragen, Konzepte, Perspektiven. R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie* (S. 3-48). Weinheim: Beltz.
- Moosbrugger, H. & A. Kelava (2008). *Testtheorie und Fragebogenkonstruktion*. Heidelberg: Springer.
- Moosbrugger, H. (2002). *Lineare Modelle*. Bern: Hans Huber.
- Moosbrugger, H. & Hartig, J. (2002). Factor analysis in personality research: Some artefacts and their consequences for psychological assessment. *Psychologische Beiträge*, 44, 136-158.
- Moosbrugger, H. & Schermelleh-Engel, K. (2008). Exploratorische (EFA) und konfirmatorische (CFA) Faktorenanalyse. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 308-324). Heidelberg: Springer.
- Mos, L. P., Wardell, D. & Royce, J. R. (1974). A factor analysis of some measures of cognitive style. *Multivariate Behavioral Research*, 9, 47-58.
- Moschner, B. (2001). Selbstkonzept. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 629-635). Weinheim: BeltzPVU.
- Moschner, B. & Gruber, H. (2005a). Which dimensions should the construct 'epistemological beliefs' comprise? *Präsentation auf der 11. Konferenzen der European Association für Learning and Instruction (EARLI) in Nicosia (Zypern)*.
- Moschner, B. & Gruber, H. (2005b). Weiterentwicklung eines Instruments zur Erfassung epistemologischer Überzeugungen. *Präsentation auf der 10. Tagung der Fachgruppe Pädagogische Psychologie der Deutschen Gesellschaft für Psychologie, Martin-Luther-Universität Halle-Wittenberg*.
- Moschner, B. & Gruber, H. (im Druck). *Fragebogen zur Erfassung epistemologischer Überzeugungen*. In A. Anschütz, H. Gruber & B. Moschner (Hrsg.), *Wissen und Lernen - in der Sicht von Lernenden und Lehrenden*. Münster: Waxmann.
- Moschner, B., Gruber, H. & die Studienstiftungsarbeitsgruppe EPI (2005). *Epistemologische Überzeugungen. Forschungsbericht Nr. 18*. Lehrstuhl Pädagogik für Lehr-Lern-Forschung und Medienpädagogik, Universität Regensburg.
- Moustaki, I. (2000). A Latent Variable Model for Ordinal Variables. *Applied Psychological Measurement*, 24, 211-223.
- Muis, K. R., Bendixen, L. D. & Haerle, F. C. (2006). Domain-general and domain-specificity in personal epistemology research: Philosophical and empirical reflections in the development of a theoretical framework. *Educational Psychology Review*, 18, 3-54.

- Mummendey, H. D. (1999). *Die Fragebogen-Methode. Grundlagen und Anwendungen in Persönlichkeits-, Einstellungs- und Selbstkonzeptforschung*. Göttingen: Hogrefe.
- Mummendey, H. D. (1990). *Psychologie der Selbstdarstellung*. Göttingen: Hogrefe.
- Mundform, D. J., Shaw, D. G. & Ke, T. L. (2005). Minimum sample size recommendations for conducting factor analysis. *International Journal of Testing*, 5, 159-168.
- Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Nunnally, J. C., Bernstein, I. (1993). *Psychometric theory*. Boston: MacGraw Hill.
- Nye, C. D., Newman, D. & Joseph, D. L. (2010). Never say "Always"? Extreme item wording effects on scalar invariance and item response curves. *Organizational Research Methods*, 13, 806-830.
- Olafson, L. & Schraw, G. (2006). Teachers' beliefs and practices within and across domains. *International Journal of Educational Research*, 45, 71-84.
- Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological Methods*, 5(3), 343-355.
- Osgood, C.E., Suci, G.J. & Tannenbaum, D.H. (1957). *The Measurement of Meaning*. Urbana: University of Illinois Press.
- Pai, Y. (1990). *Cultural Foundations of Education*. Merrill: Columbus.
- Paulsen, M. B. & Wells, C. T. (1998). Domain differences in the epistemological beliefs of college students. *Research in Higher Education*, 39, 365-384.
- Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge, England: Cambridge University Press.
- Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years: A scheme*. New York: Holt, Rinehart and Winston.
- Petermann, F. & Noack, H. (1984). Nicht-reaktive Meßverfahren. In E. Roth (Hrsg.), *Sozialwissenschaftliche Methoden* (S. 451-470). München: Oldenbourg.
- Phan, H. P. (2006). Examination of student learning approaches, reflective thinking and epistemological beliefs: A latent variable approach. *Electronic Journal of Research in Educational Psychology*, 10, 577-610.
- Phan, H. P. (2008). Multiple regression analysis of epistemological beliefs, learning approaches, and self-regulated learning. *Electronic Journal of Research in Educational Psychology*, 14, 157-184.
- Piaget, J. (1955). *The child's construction of reality*. London: Routledge.

- Pike, K. L. (1954). Emic and etic standpoints for the description of behavior. In K. L. Pike (Hrsg.), *Language in relation to a unified theory of the structure of human behavior* (S. 8–28). Glendale: Summer Institute of Linguistics.
- Pintrich, P.R., Smith, D., Garcia, T. & MacKeachie, W. (1991). *The motivated strategies for learning questionnaire (MSLQ)*. Ann Arbor: University of Michigan.
- Pintrich, P.R., Wolters, C., & Baxter, G. (2000). Assessing metacognition and self-regulated learning. In G. Schraw & J. C. Impara (Hrsg.), *Issues in the measurement of metacognition* (S. 43-97). Lincoln: Buros Institute of Mental Measurements.
- Podsakoff, P. M., MacKenzie, S. B., Lee J.-Y. & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88, 879-903.
- Preacher, K. J. & MacCallum, R. J. (2003). Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics*, 2, 13-43.
- Priemer, B. (2006). Deutschsprachige Verfahren der Erfassung von epistemologischen Überzeugungen. *Zeitschrift für Didaktik der Naturwissenschaften*, 12, 159-175.
- Qian, G. & Alvermann, D. (1995). Role of epistemological beliefs and learned helplessness in secondary school students' learning science concepts from text. *Journal of Educational Psychology*, 87, 282-292.
- Raykov, T. (1997). Scale reliability, Cronbach's alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research*, 32, 329-353.
- Raykov, T. (2001a). Bias of Cronbach's alpha for fixed congeneric measures with correlated errors. *Applied Psychological Measurement*, 25, 69-76.
- Raykov, T. (2001b). Estimation of congeneric scale reliability using covariance structure analysis with nonlinear constraints. *British Journal of Mathematical and Statistical Psychology*, 55, 145-158.
- Renkl, A. (2009). *Wissenserwerb*. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 4-26). Heidelberg: Springer.
- Richter, T. (2007). *Epistemologische Einschätzungen beim Textverstehen*. Abgerufen am 13.06.2007. Verfügbar unter <http://psydok.sulb.uni-saarland.de/volltexte/2007/910/>.
- Royce, J. R. (1959). The search for meaning. *American Scientist*, 47, 515–535.
- Royce, J. R. (1978). Three ways of knowing and the scientific world view. *Methodology and Science*, 11, 146–164.

- Royce, J. R. & Mos, L. P. (1980). *Manual: Psycho-Epistemological Profile*. Center for advanced study in theoretical psychology, University of Alberta, Canada.
- Royce, J. R. & Smith, W. A. S. (1964). A note of the development of the Psycho-epistemological Profile (PEP). *Psychological Reports*, 14, 297–298.
- Rozendaal, J. S., de Brabander, C. J. & Minnaert, A. (2001). Boundaries and dimensionality of epistemological beliefs. In C. J. de Brabander (Hrsg.), *Personal epistemology: A search for conceptual clarification*. Symposium auf der 9. Tagung der European Association for Research on Learning and Instruction, Fribourg, Schweiz.
- Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. Spiro, B. Bruce & W. Brewer (Hrsg.), *Theoretical issues in reading comprehension* (S. 33-58). Hillsdale, NJ: Erlbaum.
- Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in Personality and Social Psychology Bulletin. *Personality and Social Psychology Bulletin*, 28, 1629-1646.
- Ryan, M. P. (1984). Monitoring text comprehension: Individual differences in epistemological standards. *Journal of Educational Psychology*, 76, 248-258.
- Salomon, G. (1997). *Distributed cognitions: Psychological and educational considerations*. Cambridge: Cambridge University Press.
- Schermelleh-Engel, K., Keith, N., Moosbrugger, H. & Hodapp, V. (2004). Decomposing person and occasion-specific effects: An extension of latent state-trait theory to hierarchical LST models. *Psychological Methods*, 9, 198-219.
- Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. (2003). Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research - Online*, 8, 23-74.
- Schermelleh-Engel, K. & Schweizer, K. (2008). Multitrait-Multimethod-Analysen. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 326-341). Heidelberg: Springer.
- Schermelleh-Engel, K. & Wagner, C. (2008). Methoden der Reliabilitätsbestimmung. In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 114-133). Heidelberg: Springer.
- Schmid, J., & Leiman, J. N. (1957). The development of hierarchical factor solutions. *Psychometrika*, 22, 53-61.

- Schoenfeld, A. (1985). Metacognitive and epistemological issues in mathematical understanding. In E. A. Silver (Hrsg.), *Teaching and learning mathematical problem solving: Multiple research perspectives* (S. 361-380). Hillsdale, NJ: Erlbaum.
- Schommer-Aikins, M. (2004). Explaining the epistemological belief system: Introducing the embedded systemic model and coordinated research approach. *Educational Psychologist*, 39, 19-29.
- Schommer-Aikins, M., Brookhart, S., Hutter, R. & Mau, W.-C. (2000). Understanding middle students' beliefs about knowledge and learning using a multidimensional paradigm. *The Journal of Educational Research*, 94, 120-127.
- Schommer-Aikins, M., Duell, O. K., & Barker, S. (2003). Epistemological beliefs across domains using Biglan's classification of academic disciplines. *Research in Higher Education*, 44, 347-366.
- Schommer-Aikins, M. (2002). An evolving theoretical framework for an epistemological belief system. In B. K. Hofer & P. R. Pintrich (eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (S. 103-118). Mahwah: Lawrence Erlbaum.
- Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology*, 82, 498-504.
- Schommer, M. (1993a). Epistemological development and academic performance among secondary students. *Journal of Educational Psychology*, 85, 406-411.
- Schommer, M. (1993b). Comparisons of beliefs about the nature of knowledge and learning among postsecondary students. *Research in Higher Education*, 34, 355-369.
- Schommer, M. (1994). Synthesizing epistemological belief research: Tentative understandings and provocative confusions. *Educational Psychology Review*, 6, 293-319.
- Schommer, M. & Walker, K. (1995). Are epistemological beliefs similar across domains? *Journal of Educational Psychology*, 87, 424-432.
- Schommer, M., Crouse, A. & Rhodes, N. (1992). Epistemological beliefs and mathematical text comprehension: Believing it is simple does not make it so. *Journal of Educational Psychology*, 84, 435-443.
- Schraw, G. & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, 7, 351-371.
- Schraw, G., Bendixen, L. B. & Dunkle, M. E. (2002). *Development and validation of the epistemic belief inventory (EBI)*. In B. K. Hofer & P.R. Pintrich (Hrsg), *Personal Episte-*

- mology. The psychology of beliefs about knowledge and knowing.* Mahwah: Lawrence Erlbaum.
- Schriesheim, C. A., Eisenbach, R. J. & Hill, K. D. (1991). The effect of negation and polar opposite item reversals on questionnaire reliability and validity: An experimental investigation. *Educational and Psychological Measurement*, 51, 67-78.
- Schulz, H. (1993). *Physik mit Bleistift. Einführung in die Rechenmethoden der Naturwissenschaften.* Heidelberg: Springer.
- Schwab F. (2004). *Evolution und Emotion.* Stuttgart: Kohlhammer.
- Schweizer, K., Moosbrugger, H. & Schermelleh-Engel, K. (2003). Models for hierarchical structures in differential psychology. *Methods of Psychological Research Online*, 8, 159-180.
- Schweizer, K. & Schreiner, M. (2009). Avoiding the effect of item wording by means of bipolar instead of unipolar items: An application to social optimism. *European Journal of Personality*, 24, 137-150.
- Schwitzgebel, E. (2008). *Belief.* In Edward N. Zalta (Hrsg.), *The Stanford Encyclopedia of Philosophy* (Winter 2008 Edition). Abgerufen am 27.12.2009. Verfügbar unter <http://plato.stanford.edu/archives/win2008/entries/belief/>.
- Shadish, W. R, Cook, T. D. & Campbell, D. T. (2002). *Experimental and quasi-experimental design for generalized causal inference.* Boston: Houghton-Mifflin.
- Silvia, E. S. & Mac Callum, R. C. (1988). Some factors affecting the success of specification searches in covariance structure modeling. *Multivariate Behavioral Research*, 23, 297-326.
- Smyth, M. M. (2001). Certainty and uncertainty sciences: Marking the boundaries of psychology in introductory textbooks. *Social Studies of Science*, 31, 389–416.
- Snider, J. G. & Osgood, C. E. (1969). *Semantic differential technique. A sourcebook.* Chicago: Aldine.
- Sodian, B. (2008). Entwicklung des Denkens. In R. Oerter & L. Montada (Hrsg.), *Entwicklungspsychologie* (S. 3-48). Weinheim: Beltz.
- Sörbom, D. (1989). Model modification. *Psychometrika*, 54, 371-384.
- Spielberger, C. D. (1972). Anxiety as an emotional state. In C. D. Spielberg (Hrsg.), *Anxiety: Current trends in theory and research* (Vol. 1, S. 23-49). New York: Academic Press.
- Stahl, E. & Bromme, R. (2007). The CAEB. An instrument for measuring connotative aspects of epistemological beliefs. *Learning and Instruction*, 17, 773-785.

- Sijtsma, K. (2012). Psychological measurement between physics and statistics. *Theory and Psychology*, 22, 786-809.
- Steenkamp, J.-B. E. M. & Van Trijp, H. C. M. (1991). The use of LISREL in validating marketing constructs. *International Journal of Research in Marketing*, 8, 283-299.
- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25, 173–180.
- Stelzl, I. (1986). Changing a causal hypothesis without changing the fit: Some rules for generating equivalent LISREL models. *Multivariate Behavioral Research*, 21, 309-331.
- Stern, W. (1911). *Die differentielle Psychologie in ihren methodischen Grundlagen*. Leipzig: Barth.
- Steup, M. (2008). *Epistemology*. In Edward N. Zalta (Hrsg.)n The Stanford Encyclopedia of Philosophy (Winter 2008 Edition). Abgerufen am 26.12.2009. Verfügbar unter <http://plato.stanford.edu/archives/win2008/entries/epistemology>.
- Stewart, D. J. (1981). The application and missapplication of factor analysis in marketing research. *Journal of Marketing Research*, 18, 51-62.
- Steyer, R. (2004). Was wollen und was können wir durch empirische Kausalforschung erfahren? In E. Erdfelder & J. Funke (Hrsg.), *Allgemeine Psychologie und deduktivistische Methodologie*. Göttingen: Vandenhoeck und Ruprecht.
- Steyer, R., Ferring, D. & Schmitt, M. J. (1992). States and traits in psychological assessment. *European Journal of Psychological Assessment*, 8, 79–98.
- Steyer, R., Schmitt, M. & Eid, M. (1999). Latent state-trait theory and research in personality and individual differences. *European Journal of Personality*, 13, 389-408.
- Stodolsky, S. S., Salk, S. & Glaessner, B. (1991). Student views about learning math and social studies. *American Educational Research Journal*, 28, 89–116.
- Sudman, S., Bradburn, N. & Schwarz, N. (1996). *Thinking about answers. The application of cognitive processes to survey methodology*. San Francisco: Jossey-Bass.
- Tataryn, D. J., Wood, J. M. & Gorsuch, R. L. (1999). Setting the value of k in PROMAX: a Monte-Carlo study. *Educational & Psychological Measurement*, 59, 384-391.
- Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago: University of Chicago Press.
- Tobin, K. & McRobbie, C. (1997). Beliefs about the nature of science and the enacted science curriculum. *Science & Education*, 6, 355–371.
- Tourangeau, R., Rips, L. J. & Rasinski, K. (2000). *The psychology of survey response*. Cambridge: Cambridge University Press.

- Trafimow, D. & Rice, S. (2009). What if social scientists had reviewed great scientific works of the past? *Perspective on Psychological Science*, 4, 65-78.
- Trautwein, U. & Lüdtke, O. (2007). Predicting global and topic-specific certainty beliefs: Domain-specificity and the role of the academic environment. *British Journal of Educational Psychology*, 77, 907-934.
- Triandis, H. C. (1987). Collectivism v. individualism: A reconceptualization of a basic concept in cross-cultural social psychology. In G. K. Verma & C. Bagley (Hrsg.), *Cross cultural studies of personality, attitudes and cognition* (S. 60-95). London: Macmillan.
- Triandis, H. C. (1995). *Individualism and collectivism*. Boulder: Westview.
- Urhahne, D. (2006). Die Bedeutung domänenspezifischer epistemologischer Überzeugungen für Motivation, Selbstkonzept und Lernstrategien von Studierenden. *Zeitschrift für Pädagogische Psychologie*, 20, 189-198.
- Urhahne, D. & Hopf, M. (2004). Epistemologische Überzeugungen in den Naturwissenschaften und ihre Zusammenhänge mit Motivation, Selbstkonzept und Lernstrategien. *Zeitschrift für Didaktik der Naturwissenschaften*, 10, 71-87.
- Valsiner, J. (1986). *The individual subject and scientific psychology*. New York: Plenum Press.
- Valsiner J. (1986). Different perspectives on individual-based generalizations in psychology. In J. Valsiner (Hrsg.), *The individual subject and scientific psychology* (S. 391-404). New York: Plenum Press.
- Van Prooijen, J.-W. & Van der Kloot, W. A. (2001). Confirmatory analysis of exploratively obtained factor structures. *Educational and Psychological Measurement*, 61, 777-792.
- Vandenberg, R. J. & Lance, C. E. (2000). A review of and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4-69.
- Wason, P. C. (1966). Reasoning. In B. M. Foss (Hrsg.), *New horizons in psychology* (S. 135-151). Harmondsworth: Penguin.
- Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20, 273-281.
- Weber, R. & Crocker, J. (1993). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, 45, 961-977.
- Weinstein, C. E. (1988). Assessment and training of student learning strategies. In R. R. Schmeck (Hrsg.), *Learning strategies and learning styles*, (S. 291-316). New York: Plenum Press.

- Weise, G. (1975). *Psychologische Leistungstests*. Göttingen: Hogrefe.
- Wellman, H. (1990). *The child's theory of mind*. Cambridge: MIT press.
- West, S. G., Finch, J. F., & Curran, P. (1995). Structural equation modeling with nonnormal variables: Problems and remedies. In R. H. Hoyle (Hrsg.), *Structural equation modeling: Concepts, issues, and applications* (S. 65-75). Thousand Oaks: Sage.
- Westermann, R. (2000). *Wissenschaftstheorie und Experimentalmethodik*. Göttingen: Hogrefe.
- Widaman, K. F. (1993). Common factor analysis versus principal component analysis: Differential bias in representing model parameters? *Multivariate Behavioral Research*, 28, 263-311.
- Wild, K.-P. (2005). Individuelle Lernstrategien von Studierenden. Konsequenzen für die Hochschuldidaktik und die Hochschullehre. *Beiträge zur Lehrerbildung*, 23, 191- 206.
- Wild, K.-P. & Schiefele, U. (1994). Lernstrategien im Studium. Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 15, 185-200.
- Wild, K.-P., Schiefele, U. & Winteler, A. (2009). *LIST: Inventar zur Erfassung von Lernstrategien im Studium*. Abgerufen am 06.09.2009. Verfügbar unter <http://www-campus.uni-r.de/edu1/LIST0/list-home>.
- Willis, G. B. (2005). *Cognitive interviewing. A tool for improving questionnaire design*. Thousand Oaks: Sage Publications.
- Wittwer, J., Bromme, R. & Jucks, R. (2004). Kann man dem Internet trauen, wenn es um die Gesundheit geht? Die Glaubwürdigkeitsbeurteilung medizinischer Fachinformationen im Internet durch Laien. *Zeitschrift für Medienpsychologie*, 2, 48-56.
- Wolf, H.-G. & Preising, K. (2005). Exploring item and higher-order factor structure with the Schmid-Leiman solution: Syntax codes for SAS and SPSS. *Behavior Research Methods*, 37, 48-58.
- Wood, P. & Kardash, C. A. (2002). Critical elements in the design and analysis of studies of epistemology. In B. K. Hofer & P.R. Pintrich (Hrsg), *Personal Epistemology. The psychology of beliefs about knowledge and knowing* (S. 231-260). Mahwah: Lawrence Erlbaum.
- Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment*, 28, 186-191.

- Yung, Y-F., Thissen, D. & McLeod, L. D. (1999). Relationship between higher-order factor models and the hierarchical factor model. *Psychometrika*, 64, 113-128.
- Zimbardo, P. G. (1995). *Psychologie*. Heidelberg: Springer.
- Zimmerman, D. W. (1972). Test reliability and the Kuder-Richardson formula: Derivation from probability theory. *Educational and Psychological Measurement*, 32, 939-9.

8. Anhang

8.1 Anhang A: Epistemological Belief Questionnaire von Schommer (1990)

Items der Schommer-Fragebogens:

Item Nr.	Item
1	Wenn in einem Lehrtext ein schwer verständlicher Begriff gebraucht wird, ist es am besten, diesen zunächst einmal zu klären, bevor man weiter liest.
2	Ein Satz hat wenig Bedeutung, solange man den Kontext, in dem er gesprochen oder geschrieben wurde, nicht kennt.
3	Ein guter Student zu sein, bedeutet, auch Fakten auswendig zu lernen.
4	Ein Kurs in Lern- und Studiertechniken wäre hilfreich für mich.
5	Wie viel man lernt, hängt in erster Linie von den Dozenten ab.
6	Man kann fast alles glauben, was man für das Studium liest.
7	Ich wundere mich oft, wie viel meine Dozenten wissen.
8	Unterschiede in der Lernfähigkeit sind angeboren.
9	Es ärgert mich, wenn der Dozent nicht klar sagt, was er zu einem Problem denkt.
10	Erfolgreiche Studenten verstehen Lerninhalte schnell.
11	Ein guter Dozent muss dafür sorgen, dass die Studenten nicht von der "richtigen Spur" abkommen.
12	Wenn Wissenschaftler hart arbeiten, können sie über fast alles die Wahrheit herausfinden.
13	Leute, die Autoritäten anzweifeln, überschätzen sich.
14	Ich versuche, möglichst weitgehend die Lerninhalte verschiedener Stunden und verschiedener Lernveranstaltungen zu integrieren.
15	Die erfolgreichsten Leute haben entdeckt, wie sie ihre Lernfähigkeit verbessern können.
16	Die Dozenten machen die Dinge komplizierter als sie sind.
17	Der wichtigste Aspekt wissenschaftlichen Arbeitens ist präzise Messung und Sorgfalt.
18	Einen Text durcharbeiten heißt für mich weniger einzelne Fakten zu lernen als die Hauptideen zu erfassen.
19	Die Dozenten sollten wissen, was die beste Lehrmethode ist.
20	Ein schwieriges Kapitel immer und immer wieder zu lesen, hilft wenig, es zu verstehen.
21	Wissenschaftler können letztendlich zur Wahrheit kommen.
22	Man weiß nie, was ein Buch besagt, solange man die Absichten eines Autors nicht kennt.
23	Der wichtigste Aspekt wissenschaftlichen Arbeitens sind originelle Ideen.
24	Wenn ich die Zeit finde, einen Lehrtext noch einmal zu lesen, lerne ich beim zweiten Mal viel mehr.
25	Es liegt am jeweiligen Studenten, wie viel er aus einem Lehrbuch lernt.
26	Genialität hat mehr mit harter Arbeit als mit Intelligenz zu tun.
27	Es macht mir Spaß, über Dinge nachzudenken, über die sich Fachleute nicht einig sein können.
28	Jeder muss lernen, wie man lernt.
29	Wenn man etwas überhaupt verstehen kann, erscheint dies einem schon beim ers-

Item Nr.	Item
	ten Mal als sinnvoll.
30	Das einzig Gewisse ist die Ungewissheit.
31	Wenn man an der Universität erfolgreich sein will, stellt man am besten nicht allzu viele Fragen
32	Weisheit heißt nicht, die Antworten zu kennen, sondern zu wissen, wie man sie findet.
33	Die meisten Wörter haben eine klare Bedeutung.
34	Wahrheit ändert sich nicht.
35	Wenn man nach der Lektüre eines Textes zwar Einzelheiten vergessen hat, aber zu neuen Ideen gekommen ist, so ist dies ein Lernerfolg.
36	Bei schwierigen Entscheidungen würde ich es am liebsten haben, wenn jemand mir sagen könnte, was richtig ist.
37	Definitionen Wort für Wort auswendig zu lernen, ist oft nötig, um eine Prüfung gut zu bestehen.
38	Wenn ich lerne, präge ich mir jeweils spezifische Fakten ein.
39	Wenn man etwas nicht gleich versteht, sollte man erst recht dranbleiben.
40	Manchmal muss man die Antworten des Dozenten akzeptieren, auch wenn man sie nicht versteht.
41	Wenn die Dozenten sich mehr auf Fakten als auf Theoretisieren stützen würden, würde man mehr von ihren Veranstaltungen profitieren.
42	Ich mag Filme mit offenem Ende nicht.
43	Vorankommen im Studium heißt viel arbeiten.
44	Es ist Zeitverschwendung, sich mit Problemen zu beschäftigen, bei denen man nicht zu einer klaren und eindeutigen Antwort kommen kann.
45	Man sollte Lehrtexte immer hinterfragen.
46	Auch der Rat von Fachleuten ist oft anzuzweifeln.
47	Einige werden als gute Lerner geboren, andere haben diesbezüglich nur begrenzte Möglichkeiten.
48	Nichts außer dem Tod ist sicher.
49	Wirklich gute Studenten müssen nicht viel arbeiten, um erfolgreich ihr Studium zu bestreiten.
50	Es zahlt sich nur für die klügsten Studenten aus, sich intensiv längere Zeit mit einem Problem zu beschäftigen.
51	Wenn man zu intensiv versucht, ein Problem zu verstehen, ist man am Ende nur verwirrt.
52	Man lernt fast alles, was man aus einem Lehrtext lernen kann, beim ersten Lesen.
53	Ich verstehe schwierige Sachverhalte dann, wenn ich alle Störungen von außen abschirme und mich wirklich konzentriere.
54	Eine gute Art, einen Lehrtext zu verstehen, ist es, die Informationen entsprechend dem eigenen Vorwissen neu zu ordnen.
55	Schüler, die in der Schule durchschnittlich sind, werden im ganzen Leben "Durchschnitt" bleiben.
56	"Ein geordneter Kopf ist ein leerer Kopf".
57	Um ein Experte auf einem Gebiet zu werden, muss man ein besonderes Talent dafür mitbringen.
58	Ich schätze es, wenn Dozenten ihre Lehrveranstaltungen minutiös planen und sich dann an ihren Plan halten.
59	Das Beste an der Wissenschaft ist, dass sie zeigt, dass es auf die meisten Probleme eine richtige Antwort gibt.
60	Lernen ist ein langsamer Prozess des Wissensaufbaus.

Item Nr.	Item
61	Derzeit gültige Fakten sind meist schon bald veraltet.
62	Bücher zur Selbsthilfe sind von geringem Nutzen.
63	Es verwirrt nur, wenn man die Inhalte eines Lehrtextes mit dem in Verbindung bringt, was man bereits weiß.

Zuordnung zu den Subsets:

Skala Nr.	Skala	Anzahl Items	Items
S1	Success is unrelated to hard work	4	26-, 32-, 43-, 49+
S2	Avoid ambiguity	5	9+, 27-, 41+, 42+, 44+
S3	Seek single answers	11	2-, 11+, 16+, 17+, 19+, 22-, 23-, 33+, 56-, 58+, 59+
S4	Avoid Integration	8	3+, 14-, 18-, 35+, 37+, 38+, 54-, 63+
S5	Depend on authority	4	1-, 5+, 36+, 40+
S6	Ability to learn is innate	4	8+, 47+, 55+, 57+
S7	Don't criticize authority	6	6+, 7-, 13+, 31+, 45-, 46-
S8	Knowledge is certain	6	12+, 21+, 30-, 34+, 48-, 61-
S9	Learning is quick	5	10+, 29+, 39-, 50+, 60-
S10	Can't learn how to learn	5	4-, 15-, 25-, 28-, 62+
S11	Learn first time	3	20+, 24-, 52+
S12	Concentrated effort is a waste of time	2	51+, 53-

Zuordnung der Subsets zu den Faktoren:

Faktor	Subsets	Subset Nr.
Simple knowledge	Seek single answers Avoid integration	S3, S4
Certain knowledge	Avoid ambiguity Knowledge is certain	S2, S8
Omniscient authority	Depend on authority Don't criticize authority	S5, S7
Innate ability	Success is unrelated to hard work Ability to learn is innate Can't learn how to learn	S1, S6, S10
Quick learning	Learning is quick Learn first time Concentrated effort is a waste of time	S9, S11, S12

8.2 Anhang B: EPI-Fragebogen

Items des EPI-Fragebogens (Items, die rekodiert werden müssen, sind hinter der Itemnummer mit einem R gekennzeichnet.)

1. Wissen

Reflexive Natur des Wissens

Item Nr.	Item
7	Wissen entwickelt sich weiter, wenn man sich kritisch damit auseinandersetzt.
27	Durch neue Erkenntnisse wird häufig das bisherige Wissen in Frage gestellt.
28	Durch die Beschäftigung mit neuen Wissensinhalten erscheinen bekannte Wissensinhalte oft in einem ganz anderen Licht.
48	Nach intensivem Nachdenken sieht man Probleme oft mit anderen Augen.
51	Die Beurteilung von Wissen verändert sich mit neuen Erfahrungen.

Absolutes Wissen

Item Nr.	Item
1	Wenn man etwas in einem Lehrbuch liest, dann kann man auch sicher sein, dass es wahr ist.
4	Alle Expertinnen/Experten in einem Fach haben das gleiche Grundverständnis.
9	Man kann fast alles glauben, was man für das Studium liest.
13	Es gibt wissenschaftliche Erkenntnisse, die immer gültig sein werden.
15	Alle Dozentinnen/Dozenten im gleichen Fach kommen wahrscheinlich zu den gleichen Antworten auf fachliche Fragen.
16	Wenn Probleme besonders strittig sind, wendet man sich am besten an Expertinnen/Experten.
17	Wissenschaftlerinnen/Wissenschaftler können letztendlich zur Wahrheit kommen.
29	Die Antworten von Dozentinnen/Dozenten müssen als wahr akzeptiert werden.
34	In der Wissenschaft werden verschiedene Phänomene der Welt objektiv erklärt.
37	Die Kernaussagen in der Wissenschaft sind in hohem Maße objektiv.
42	Dozentinnen/Dozenten haben fast immer Recht.
44	Es gibt unumstößliche Wahrheiten.
45	In der Wissenschaft gibt es einen festen Kern von Wissen.
49	Die wissenschaftliche Forschung zeigt, dass es auf die meisten Probleme eine richtige Antwort gibt.

Soziale Komponente des Wissens

Item Nr.	Item
20	Nur veröffentlichte Erkenntnisse können als Wissen gelten.
22R	Es gibt Wissen, das nur einer Person bekannt ist.
25R	Ein Gelehrter besitzt Wissen, auch wenn er auf einer einsamen Insel lebt.
40R	Wissen existiert auch, wenn man es für sich behält.
41	Nur wenn jemand auch in der Lage ist, sein Wissen anderen mitzuteilen, kann man sicher sein, dass das Wissen auch vorhanden ist.
52	Wer sein Wissen nicht zeigt, weiß auch nichts.

Value of knowledge

Item Nr.	Item
12	Wissenschaftliche Forschung sollte immer eine praktische Relevanz haben.
21	Man sollte vor allem Fragen untersuchen, die zur Lösung praktischer Probleme beitragen.
24	Wissenschaftliche Erkenntnisse, die keinen praktischen Nutzen haben, sind wertlos.
38	Wissen ohne gesellschaftlichen Nutzen ist wertlos.
53	Der gesellschaftliche Nutzen bestimmt den Wert neuer Erkenntnisse.

2. Lernen**Geschlechtsspezifische Wissenszugänge**

Item Nr.	Item
5R	Die Aufnahmefähigkeit von Männern und Frauen ist gleich.
8	Die Fähigkeit zum Auswendiglernen unterscheidet sich zwischen Männern und Frauen.
14R	Frauen und Männer wissen gleich viel.
18R	In Naturwissenschaften sind Frauen und Männer gleich gut begabt.
19R	Die Lerngeschwindigkeit ist bei Frauen und Männern gleich.
23	Männer und Frauen analysieren Texte unterschiedlich.
30	Manche Wissensbereiche erschließen sich Frauen anders als Männern.
36	Frauen und Männer bevorzugen unterschiedliche Zugänge zum Wissen.
43	Männer und Frauen haben unterschiedliche Begabungen für Sprachen.
46	Frauen und Männer deuten Wissen auf unterschiedliche Weise.

Kulturspezifische Wissenszugänge

Item Nr.	Item
11	Wie Wissen vermittelt wird, ist von Kultur zu Kultur unterschiedlich.
32	Welches Wissen wichtig ist, hängt von der Kultur ab.
35	Der Wissenserwerb wird stark von der Kultur beeinflusst, in der man aufwächst.
39	Menschen aus verschiedenen Kulturen erwerben Wissen auf unterschiedliche Weise.
10R	Kulturunterschiede haben keinen besonderen Einfluss auf den Wissenserwerb.
26R	Menschen lernen auf der ganzen Welt gleich.
50	Die Einstellung einer Person zum Lernen wird durch ihre Kultur beeinflusst.

Lernen lernen

Item Nr.	Item
2	Schwache Lerner können sich noch so anstrengen, sie werden niemals zu guten Lernern.
3	Es ist kaum möglich, das Lernen zu lernen.
6R	Alle Studierenden sollten Kurse über Lerntechniken besuchen.
31R	Jeder muss lernen, wie man lernt.
33R	Der gezielte Einsatz von Lerntechniken führt zu höheren Leistungen.
47R	Man kann das Lernen lernen.

8.3 Anhang C: Fragebogen zur Erfassung epistemologischer Überzeugungen (FEE)

Sicherheit von Wissen (Cronbachs $\alpha = .78$)

Item Nr.	Item	<i>M</i>	<i>SD</i>	r_{it}
13	Es gibt wissenschaftliche Erkenntnisse, die immer gültig sein werden.	4.10	1.61	.54
17	Wissenschaftlerinnen/Wissenschaftler können letztendlich zur Wahrheit kommen.	3.16	1.37	.45
34	In der Wissenschaft werden verschiedene Phänomene der Welt objektiv erklärt.	3.95	1.20	.46
37	Die Kernaussagen in der Wissenschaft sind in hohem Maße objektiv.	3.70	1.18	.48
44	Es gibt unumstößliche Wahrheiten.	4.01	1.72	.56
45	In der Wissenschaft gibt es einen festen Kern von Wissen.	3.88	1.25	.55
49	Die wissenschaftliche Forschung zeigt, dass es auf die meisten Probleme eine richtige Antwort gibt.	3.27	1.21	.47

Umgang mit Autoritäten (Cronbachs $\alpha = .71$)

Item Nr.	Item	<i>M</i>	<i>SD</i>	r_{it}
4	Alle Expertinnen/Experten in einem Fach haben das gleiche Grundverständnis.	2.30	1.20	.42
9	Man kann fast alles glauben, was man für das Studium liest.	2.50	1.22	.48
15	Alle Dozentinnen/Dozenten im gleichen Fach kommen wahrscheinlich zu den gleichen Antworten auf fachliche Fragen.	2.35	1.20	.47
16	Wenn Probleme besonders strittig sind, wendet man sich am besten an Expertinnen/Experten.	2.08	1.21	.49
29	Die Antworten von Dozentinnen/Dozenten müssen als wahr akzeptiert werden.	2.45	1.24	.48
42	Dozentinnen/Dozenten haben fast immer Recht.	2.30	1.20	.42

Reflexive Natur von Wissen (Cronbachs $\alpha = .79$)

Item Nr.	Item	<i>M</i>	<i>SD</i>	r_{it}
7	Wissen entwickelt sich weiter, wenn man sich kritisch damit auseinandersetzt.	5.13	.94	.49
27	Durch neue Erkenntnisse wird häufig das bisherige Wissen in Frage gestellt.	4.79	1.05	.54
28	Durch die Beschäftigung mit neuen Wissensinhalten erscheinen bekannte Wissensinhalte oft in einem ganz anderen Licht.	4.93	.96	.65
48	Nach intensivem Nachdenken sieht man Probleme oft mit anderen Augen.	4.91	.94	.55
51	Die Beurteilung von Wissen verändert sich mit neuen Erfahrungen.	4.92	.89	.60

Soziale Komponente von Wissen (Cronbachs $\alpha = .65$)

Item Nr.	Item	<i>M</i>	<i>SD</i>	r_{it}
20	Nur veröffentlichte Erkenntnisse können als Wissen gelten.	4.73	1.31	.40
25	Ein Gelehrter besitzt Wissen, auch wenn er auf einer einsamen Insel lebt.	4.87	1.29	.39
40	Wissen existiert auch, wenn man es für sich behält.	4.72	1.31	.46
41	Nur wenn jemand auch in der Lage ist, sein Wissen anderen mitzuteilen, kann man sicher sein, dass das Wissen auch vorhanden ist.	4.06	1.52	.34
52	Wer sein Wissen nicht zeigt, weiß auch nichts.	5.28	.96	.50

Wertigkeit von Wissen (Cronbachs $\alpha = .72$)

Item Nr.	Item	<i>M</i>	<i>SD</i>	r_{it}
12	Wissenschaftliche Forschung sollte immer eine praktische Relevanz haben.	4.52	1.31	.42
21	Man sollte vor allem Fragen untersuchen, die zur Lösung praktischer Probleme beitragen.	3.84	1.32	.50
24	Wissenschaftliche Erkenntnisse, die keinen praktischen Nutzen haben, sind wertlos.	2.39	1.32	.52
38	Wissen ohne gesellschaftlichen Nutzen ist wertlos.	2.78	1.34	.49
53	Der gesellschaftliche Nutzen bestimmt den Wert neuer Erkenntnisse.	3.28	1.37	.47

Geschlechtsspezifische Wissenszugänge (Cronbachs $\alpha = .83$)

Item Nr.	Item	<i>M</i>	<i>SD</i>	r_{it}
23	Männer und Frauen analysieren Texte unterschiedlich.	4.24	1.29	.65
30	Manche Wissensbereiche erschließen sich Frauen anders als Männern.	4.55	1.24	.60
36	Frauen und Männer bevorzugen unterschiedliche Zugänge zum Wissen.	4.10	1.16	.65
43	Männer und Frauen haben unterschiedliche Begabungen für Sprachen.	4.04	1.40	.57
46	Frauen und Männer deuten Wissen auf unterschiedliche Weise.	3.84	1.19	.63

Kulturspezifische Wissenszugänge (Cronbachs $\alpha = .78$)

Item Nr.	Item	<i>M</i>	<i>SD</i>	r_{it}
11	Wie Wissen vermittelt wird, ist von Kultur zu Kultur unterschiedlich.	4.71	1.06	.55
26	Menschen lernen auf der ganzen Welt gleich.	4.84	1.17	.30
32	Welches Wissen wichtig ist, hängt von der Kultur ab.	4.26	1.28	.45
35	Der Wissenserwerb wird stark von der Kultur beeinflusst, in der man aufwächst.	4.62	1.04	.62
39	Menschen aus verschiedenen Kulturen erwerben Wissen auf unterschiedliche Weise.	4.48	1.08	.65
50	Die Einstellung einer Person zum Lernen wird durch ihre Kultur beeinflusst.	4.33	1.07	.63

Lernen lernen (Cronbach's $\alpha = 0,73$)

Item Nr.	Item	<i>M</i>	<i>SD</i>	r_{it}
6	Alle Studierenden sollten Kurse über Lerntechniken besuchen.	3.95	1.49	.48
31	Jeder muss lernen, wie man lernt.	4.40	1.39	.52
33	Der gezielte Einsatz von Lerntechniken führt zu höheren Leistungen.	4.69	1.04	.57
47	Man kann das Lernen lernen.	4.92	1.09	.57

8.4 Anhang D: Inventar zur Erfassung von Lernstrategien im Studium (LIST)

Organisation

Item Nr.	Item
1	Ich fertige Tabellen, Diagramme oder Schaubilder an, um den Stoff der Veranstaltung besser strukturiert vorliegen zu haben.
14	Ich mache mir kurze Zusammenfassungen der wichtigsten Inhalte als Gedankenstütze.
25	Ich gehe meine Aufzeichnungen durch und mache mir dazu eine Gliederung mit den wichtigsten Punkten.
36	Ich versuche, den Stoff so zu ordnen, daß ich ihn mir gut einprägen kann.
47	Ich stelle mir aus Mitschrift, Skript oder Literatur kurze Zusammenfassungen mit den Hauptideen zusammen.
56	Ich unterstreiche in Texten oder Mitschriften die wichtigen Stellen.
65	Für größere Stoffmengen fertige ich eine Gliederung an, die die Struktur des Stoffs am besten wieder gibt.
72	Ich stelle wichtige Fachausdrücke und Definitionen in eigenen Listen zusammen.

Elaborieren

Item Nr.	Item
2	Ich versuche, Beziehungen zu den Inhalten verwandter Fächer bzw. Lehrveranstaltungen herzustellen.
17	Zu neuen Konzepten stelle ich mir praktische Anwendungen vor.
28	Ich versuche, neue Begriffe oder Theorien auf mir bereits bekannte Begriffe und Theorien zu beziehen.
39	Ich stelle mir manche Sachverhalte bildlich vor.
49	Ich versuche in Gedanken, das Gelernte mit dem zu verbinden, was ich schon darüber weiß.
58	Ich denke mir konkrete Beispiele zu bestimmten Lerninhalten aus.
67	Ich beziehe das, was ich lerne, auf meine eigenen Erfahrungen.
73	Ich überlege mir, ob der Lernstoff auch für mein Alltagsleben von Bedeutung ist.

Kritisches Prüfen

Item Nr.	Item
3	Ich frage mich, ob der Text, den ich gerade durcharbeite, wirklich überzeugend ist.
18	Ich prüfe, ob die in einem Text (oder in meiner Mitschrift) dargestellten Theorien, Interpretationen oder Schlussfolgerungen ausreichend belegt und begründet sind
29	Ich denke über Alternativen zu den Behauptungen oder Schlussfolgerungen in den Lerntexten nach.
40	Der Stoff, den ich gerade bearbeite, dient mir als Ausgangspunkt für die Entwicklung eigener Ideen.
50	Es ist für mich sehr reizvoll, widersprüchliche Aussagen aus verschiedenen Texten aufzuklären.
59	Ich gehe an die meisten Texte kritisch heran.
68	Ich vergleiche die Vor- und Nachteile verschiedener theoretischer Konzeptionen.
74	Das, was ich lerne, prüfe ich auch kritisch.

Wiederholen

Item Nr.	Item
4	Ich präge mir den Lernstoff von Texten durch Wiederholen ein.
19	Ich lese meine Aufzeichnungen mehrmals hintereinander durch.
30	Ich lerne Schlüsselbegriffe auswendig, um mich in der Prüfung besser an wichtige Inhaltsbereiche erinnern zu können.
41	Ich lerne eine selbst erstellte Übersicht mit den wichtigsten Fachtermini auswendig.
51	Ich lese einen Text durch und versuche, ihn mir am Ende jedes Abschnitts auswendig vorzusagen.
60	Ich lerne Regeln, Fachbegriffe oder Formeln auswendig.
69	Ich lerne den Lernstoff anhand von Skripten oder anderen Aufzeichnungen möglichst auswendig.

Metakognitive Strategien

Item Nr.	Item
5	Ich versuche, mir vorher genau zu überlegen, welche Teile eines bestimmten Themengebiets ich lernen muss und welche nicht.
6	Wenn ich einen schwierigen Text vorliegen habe, passe ich meine Lerntechnik den höheren Anforderungen an (z.B. durch langsames Lesen).
13	Wenn ich während des Lesens eines Textes nicht alles verstehe, versuche ich, die Lücken festzuhalten und den Text daraufhin noch einmal durchzugehen.
20	Ich lege im Vorhinein fest, wie weit ich mit der Durcharbeitung des Stoffs kommen möchte.
31	Vor dem Lernen eines Stoffgebiets überlege ich mir, wie ich am effektivsten vorgehen kann.
42	Ich überlege mir vorher, in welcher Reihenfolge ich den Stoff durcharbeite.
52	Ich stelle mir Fragen zum Stoff, um sicherzugehen, dass ich auch alles verstanden habe.
61	Um Wissenslücken festzustellen, rekapituliere ich die wichtigsten Inhalte, ohne meine Unterlagen zu Hilfe zu nehmen.
70	Ich bearbeite zusätzliche Aufgaben, um festzustellen, ob ich den Stoff wirklich verstanden habe.
75	Um mein eigenes Verständnis zu prüfen, erkläre ich bestimmte Teile des Lernstoffs einem Studienkollegen.
77	Wenn mir eine bestimmte Textstelle verworren und unklar erscheint, gehe ich sie noch einmal langsam durch.

Anstrengung

Item Nr.	Item
9	Wenn ich mir ein bestimmtes Pensum zum Lernen vorgenommen habe, bemühe ich mich, es auch zu schaffen.
21	Ich streng mich auch dann an, wenn mir der Stoff überhaupt nicht liegt.
32	Ich gebe nicht auf, auch wenn der Stoff sehr schwierig oder komplex ist.
43	Ich lerne auch spätabends und am Wochenende, wenn es sein muss.
53	Gewöhnlich dauert es nicht lange, bis ich mich dazu entschließe, mit dem Lernen anzufangen.
62	Vor der Prüfung nehme ich mir ausreichend Zeit, um den ganzen Stoff noch einmal durchzugehen.
71	Ich nehme mir mehr Zeit zum Lernen als die meisten meiner Studienkollegen.
76	Ich arbeite so lange, bis ich mir sicher bin, die Prüfung gut bestehen zu können.

Aufmerksamkeit

Item Nr.	Item
10	Beim Lernen merke ich, dass meine Gedanken abschweifen.
22	Es fällt mir schwer, bei der Sache zu bleiben.
33	Ich ertappe mich dabei, dass ich mit meinen Gedanken ganz woanders bin.
44	Beim Lernen bin ich unkonzentriert.
54	Wenn ich lerne, bin ich leicht abzulenken.
63	Meine Konzentration hält nicht lange an.

Zeitmanagement

Item Nr.	Item
11	Beim Lernen halte ich mich an einen bestimmten Zeitplan.
23	Ich lege bestimmte Zeiten fest, zu denen ich dann lerne.
34	Ich lege die Stunden, die ich täglich mit Lernen verbringe, durch einen Zeitplan fest.
45	Ich lege vor jeder Lernphase eine bestimmte Zeitdauer fest.

Lernumgebung

Item Nr.	Item
12	Ich lerne an einem Platz, wo ich mich gut auf den Stoff konzentrieren kann.
24	Ich gestalte meine Umgebung so, dass ich möglichst wenig vom Lernen abgelenkt werde.
35	Zum Lernen sitze ich immer am selben Platz.
46	Wenn ich lerne, Sorge ich dafür, dass ich in Ruhe arbeiten kann.
55	Mein Arbeitsplatz ist so gestaltet, dass ich alles schnell finden kann.
64	Die wichtigsten Unterlagen habe ich an meinem Arbeitsplatz griffbereit.

Lernen mit Studienkollegen

Item Nr.	Item
7	Ich bearbeite Texte oder Aufgaben zusammen mit meinen Studienkollegen.
15	Ich nehme mir Zeit, um mit Studienkollegen über den Stoff zu diskutieren.
26	Ich vergleiche meine Vorlesungsmitschriften mit denen meiner Studienkollegen.
37	Ich lasse mich von einem Studienkollegen abfragen und stelle auch ihm Fragen zum Stoff.
48	Ich nehme die Hilfe anderer in Anspruch, wenn ich ernsthafte Verständnisprobleme habe.
57	Wenn mir etwas nicht klar ist, so frage ich einen Studienkollegen um Rat.
66	Entdecke ich größere Lücken in meinen Aufzeichnungen, so wende ich mich an meine Studienkollegen.

Literatur

Item Nr.	Item
8	Ich suche nach weiterführender Literatur, wenn mir bestimmte Inhalte noch nicht ganz klar sind.
16	Wenn ich einen Fachbegriff nicht verstehe, so schlage ich in einem Wörterbuch nach.
27	Fehlende Informationen suche ich mir aus verschiedenen Quellen zusammen (z.B. Mitschriften, Bücher, Fachzeitschriften).
38	Ich ziehe zusätzlich Literatur heran, wenn meine Aufzeichnungen unvollständig sind.
