

Inhaltsverzeichnis

Veröffentlichungen im Rahmen des Dissertationsprojekts	i
Dank	iii
1. Vorsatz	1
I Theorie	
2. Entwicklung und „Validierung“ eines Testinstruments	10
2.1. Validität	11
2.1.1. Definition und grundlegende Eigenschaften	11
2.1.2. Klassische Arten der Validität	14
2.1.3. Konstruktvalidität als übergeordnetes Konzept	18
2.1.4. Argumentbasierter Ansatz der Validierung	22
2.1.5. Zum Verhältnis von Validität zu anderen Testgütekriterien	25
2.1.6. Zusammenfassung	26
2.2. Entwicklung eines Testinstruments	27
2.2.1. Konstruktirrelevante Varianz	28
2.2.2. Klassische Beurteilungsfehler	30
2.2.3. Aspekte der Testplanung	31
2.2.4. Aspekte der Testkonstruktion	33
2.2.5. Itemerprobung, Review und Überarbeitung	37
2.2.6. Besonderheiten der Fragebogenmethode	37
2.2.7. Zusammenfassung	40
3. Selbstwirksamkeitserwartungen	41
3.1. Mehrperspektivische theoretische Annäherung an das Konstrukt	42
3.1.1. Die sozial-kognitive Theorie Banduras	43
3.1.2. Exkurs: Philosophische Randnotizen	45
3.1.3. Selbstwirksamkeitserwartungen als Kontroll-Konstrukt	46

3.2.	Analyse der Definition: Eigenschaften und Struktur von Selbstwirksamkeitserwartungen	49
3.2.1.	Die Definition des Konstruktes	50
3.2.2.	Weitere Eigenschaften des Konstruktes, die sich aus der Definition und der theoretischen Einbettung ergeben	54
3.2.3.	Weitere strukturelle Eigenschaften des Konstruktes	56
3.3.	Quellen von Selbstwirksamkeitserwartungen	57
3.4.	Abgrenzung und Beziehung zu anderen Konstrukten	60
3.4.1.	Persönlichkeit	61
3.4.2.	Einstellungen (“attitudes“)	62
3.4.3.	Selbstkonzept	63
3.4.4.	Selbstwertgefühl	69
3.4.5.	Effektanzmotivation (“effectance motivation“)	70
3.4.6.	Kontrollüberzeugungen (“locus of control“)	70
3.4.7.	Handlungsergebniserwartungen („outcome expectancies“)	72
3.5.	Die Bedeutung von Selbstwirksamkeitserwartungen	73
3.6.	Zur Erfassung des Konstruktes	75
3.6.1.	Spezifitätsniveau	75
3.6.2.	Anleitung zur Konstruktion und Validierung von Selbstwirksamkeitserwartungsskalen	77
3.7.	Exkurs: Kollektive Selbstwirksamkeitserwartungen	80
3.8.	Zusammenfassung	81
4.	Von „teacher efficacy“ zu „Lehrer-Selbstwirksamkeitserwartungen“	83
4.1.	Definitionsversuche	85
4.1.1.	Definition nach Rotter	86
4.1.2.	Definition nach Bandura	87
4.1.3.	Theoriekonfusion: Wie Gibson und Dembo (1984) versuchen, „teacher efficacy“ und Banduras sozial-kognitive Theorie zusammenzubringen	88
4.1.4.	Das Modell von Tschannen-Moran, Woolfolk Hoy und Hoy (1998): Ausweg aus der Krise?	90
4.1.5.	Das Ende der begrifflichen Verwirrung: Dellinger, Bobbett, Olivier und Ellett (2008) bringen Licht ins Dunkel	93
4.1.6.	Vertiefte Auseinandersetzung mit dem Ansatz von Dellinger, Bobbett, Olivier und Ellett (2008): Auf dem Weg zu einem besseren Verständnis des Konstruktes der Lehrer-Selbstwirksamkeitserwartungen	95
4.1.7.	Das Konstrukt der Lehrer-Selbstwirksamkeitserwartungen im deutschen Sprachraum	100

4.1.8.	Zusammenfassung	101
4.2.	Abgrenzung und Beziehung zu verwandten Konstrukten	102
4.2.1.	Lehrer-Selbstkonzept/“teacher self-concept“	102
4.2.2.	Kontrollüberzeugungen von Lehrpersonen/“teacher locus of control“	102
4.3.	Zur Erfassung von Lehrer-Selbstwirksamkeitserwartungen	104
4.3.1.	Personal Teaching Efficacy Vignette Scale	106
4.3.2.	Teacher Efficacy Scale – TES	111
4.3.3.	Instrument zur Erfassung von Lehrer-Selbstwirksamkeitserwartungen nach Schmitz und Schwarzer (2000)	114
4.3.4.	Teacher Self-Efficacy Scale – TSES	116
4.3.5.	Ohio State Teacher Efficacy Scale – OSTES	117
4.3.6.	Teacher Confidence Scale – TCS	120
4.3.7.	Norwegian Teachers Self-Efficacy Scale – NTSES	120
4.3.8.	Mehrdimensionale Skala der Lehrer-Selbstwirksamkeitserwartung – MSLS	121
4.3.9.	Teachers’ Efficacy Beliefs System-Self – TEBS-Self	122
4.3.10.	Self-Efficacy Measure for Student Teachers in Competence-Based Education	123
4.3.11.	Zusammenfassung	124
4.3.12.	Ableitungen für die Instrumententwicklungen	127
4.4.	Empirische Basis zur Relevanz des Konstruktes	128
4.4.1.	Lehrervariablen	132
4.4.2.	Schülervariablen	135
4.4.3.	Zusammenfassung	136
4.5.	Quellen von Lehrer-Selbstwirksamkeitserwartungen	137
4.5.1.	Einflussfaktoren auf Lehrer-Selbstwirksamkeitserwartungen	138
4.6.	Zur Entwicklung von Lehrer-Selbstwirksamkeitserwartungen	140
4.6.1.	Ergebnisse querschnittlich angelegter Studien	140
4.6.2.	Ergebnisse längsschnittlich angelegter Studien	141
4.6.3.	Zusammenfassung und Kommentar zur Problematik der Interpretation der Studienergebnisse aufgrund der Erfassung des Konstruktes mittels defizitärer Instrumente	143
4.7.	Zur Ausprägung von Lehrer-Selbstwirksamkeitserwartungen in unterschiedlichen Gruppen	144
4.8.	Zusammenfassung	145
5.	Lehrer-Selbstwirksamkeitserwartungen in physikdidaktischen Handlungsfeldern (“science teaching self-efficacy beliefs“)	149
5.1.	Zur Definition des Konstruktes	150

5.2.	Zur Erfassung des Konstruktes	151
5.2.1.	Science Teaching Efficacy Beliefs Instrument – STEBI	154
5.2.2.	Self-Efficacy Teaching and Knowledge Instrument for Science Teachers – SETAKIST	156
5.2.3.	Self-Efficacy Beliefs about Equitable Science Teaching and Learning – SEBEST	156
5.2.4.	Selbstwirksamkeitserwartung bezogen auf das Physik-Lehren	157
5.2.5.	Zusammenfassung	158
5.3.	Abschließende Bemerkungen	158
6.	Von der „guten“ Lehrperson – Selbstwirksamkeitserwartungen im Forschungsfeld der Lehrprofessionalisierung	160
6.1.	Kompetenzorientierung in der Lehrerbildung	162
6.1.1.	Kompetenz – Annäherung an eine Definition	162
6.1.2.	Modellierung der Kompetenzen von (angehenden) Lehrpersonen	164
6.1.3.	Kompetenzentwicklung von Lehrpersonen	169
6.1.4.	Messung von Lehrerkompetenzen	170
6.2.	Kompetenzmodellierungen von (angehenden) Physiklehrkräften	172
6.2.1.	Projekte mit dem Schwerpunkt der Kompetenzmodellierung von (angehenden) Physiklehrkräften	173
6.2.2.	Physikdidaktisches Wissen	176
6.3.	Zusammenfassung	179
 II Instrumententwicklung, Forschungsfragen und Studiendesign		
7.	Theoriebasierte Instrumententwicklung – Erste Schritte	183
7.1.	Theoretische Ableitungen	184
7.1.1.	Arbeitsdefinition	184
7.1.2.	Konstruktionsregeln bez. des Itemstammes	185
7.1.3.	Itemkonstruktion unter Berücksichtigung des Bezugs zu anderen Konstrukten	188
7.1.4.	Wahl des Antwortformates	188
7.1.5.	Reflexion des formalen Status zugrunde liegender Theorien	188
7.2.	Struktur des Konstruktes	189
7.2.1.	Festlegung des Spezifitätsniveaus	189
7.2.2.	Inhalte/Handlungsfelder	190
7.2.3.	Dimensionen	192
7.2.4.	Hierarchie	192
7.2.5.	Zusammenfassung	193

7.3.	Vorgehen bei der Itemkonstruktion	194
7.4.	Testplanung	194
7.5.	Festlegung der zu validierenden Testwertinterpretation	196
7.5.1.	Intendierte Testwertinterpretation	196
7.5.2.	Zugrunde liegende Annahmen der intendierten Testwertinterpretation	197
7.5.3.	Gewählte Verfahren zur Prüfung der Gültigkeit getroffener Annahmen	199
7.5.4.	Konkurrierende Testwertinterpretationen	201
7.5.5.	Abschätzung potentieller Konsequenzen	201
7.6.	Zusammenfassung	202
8.	Forschungsfragen, Hypothesen, Studiendesign	204
8.1.	Forschungsfragen und Hypothesen	204
8.1.1.	Übergeordnete Forschungsfrage	204
8.1.2.	Teilforschungsfragen und zugehörige Hypothesen	204
8.2.	Projektdesign	209
8.2.1.	Übersicht des Projektablaufs	209
8.2.2.	Vorstellung einzelner Projektteile	210
III Empirie		
9.	Exkurs: Pilotstudie I	215
9.1.	Stichprobe	215
9.2.	Fragebogen	216
9.3.	Methodische Anmerkungen	216
9.4.	Analyseverfahren und benutzte Analysesoftware	216
9.5.	Ergebnisse	216
9.5.1.	Reliabilitätsanalyse	216
9.5.2.	Konfirmatorische Faktorenanalyse	217
9.5.3.	Korrelationsanalyse	217
9.5.4.	Varianzanalyse	219
9.6.	Interpretation der Ergebnisse und Validitätsargumente	219
9.6.1.	Reliabilitätsanalyse	219
9.6.2.	Konfirmatorische Faktorenanalyse	220
9.6.3.	Korrelationsanalyse	220
9.6.4.	Varianzanalyse	221
9.7.	Itemrevision	221
9.8.	Zusammenfassung	221

10. Pilotstudien II und III (qualitativer Natur)	223
10.1. Interviewstudie	223
10.1.1. Verfahren der Gewinnung von Interviewpartnern	223
10.1.2. Stichprobe	224
10.1.3. Methodische Anmerkungen	224
10.1.4. Datenerhebung – Interviewdurchführung	226
10.1.5. Datenaufbereitung und Auswerteverfahren	227
10.2. Expertenbefragung	228
10.2.1. Verfahren der Gewinnung von Experten	228
10.2.2. Methodische Anmerkungen – Aufbau des Fragebogens	228
10.2.3. Stichprobe	229
10.2.4. Datenaufbereitung und Auswerteverfahren	229
10.3. Exkurs: Akzeptanzbefragung bez. des Antwortformates	230
10.4. Ergebnisse der Pilotstudien II und III	230
10.4.1. Umsetzung der Arbeitsdefinition	231
10.4.2. Itemverständnis	231
10.4.3. Aktivierung kognitiver Prozesse bei der Itembeantwortung	232
10.4.4. Authentizität der Items	232
10.4.5. Exkurs: Itemschwierigkeiten	233
10.4.6. Eignung des Antwortformates	234
10.4.7. Angemessenheit der Abbildung der Handlungsfelder	237
10.4.8. Trennschärfe der Konstrukte	237
10.4.9. Exkurs: Relevanz der Handlungsfelder	237
10.5. Itemrevision	239
10.5.1. Vorgehen und Ableitungen bez. der Itemüberarbeitung	239
10.5.2. Exemplarische Illustration der Itemüberarbeitung	241
10.6. Interpretation der Ergebnisse und Validitätsargumente	242
10.6.1. Aktivierung kognitiver Prozesse und Umsetzung der Arbeitsdefinition	242
10.6.2. Itemverständnis	243
10.6.3. Relevanz und Authentizität	243
10.6.4. Eignung des Antwortformates	244
10.6.5. Trennschärfe der Konstrukte	244
10.7. Zusammenfassung	244
11. Weiterentwicklung des Instruments	245
11.1. Überarbeitung der Itemkonstruktionsregeln	245
11.2. Überarbeitung des Antwortformates	246
11.3. Zusammensetzung des Fragebogens	246
11.3.1. Demografie	246

11.3.2. Items	246
11.3.3. Zusätzliche Konstrukte	248
11.4. Instruktionen	250
11.5. Anordnung der Skalen	251
11.6. Überlegungen bez. der Tendenz zu sozialer Erwünschtheit	252
11.7. Papier-und-Bleistift versus Online-Befragung	252
12. Hauptstudie	253
12.1. Datenerhebung	254
12.1.1. Rekrutierung von Teilnehmerinnen und Teilnehmern	254
12.1.2. Durchführung der Datenerhebung	255
12.2. Datenaufbereitung	255
12.2.1. Datenselektion	255
12.3. Datenqualität	256
12.3.1. Fehlende Werte	257
12.3.2. Verteilung der Item- und Skalenwerte	257
12.4. Stichprobe und deskriptive Analysen	258
12.4.1. Detaillierte Beschreibung der befragten Kohorten	259
12.4.2. Beschreibung relevanter Subgruppen	260
12.5. Analyseverfahren und benutzte Analysesoftware	261
12.6. Methodische Anmerkungen	261
12.6.1. Anmerkungen zur klassischen Item- & Skalenanalyse	262
12.6.2. Konfirmatorische Faktorenanalyse	262
12.6.3. Raschanalyse	264
12.6.4. Korrelationsanalysen	267
12.6.5. Mittelwertvergleiche	267
12.7. Ergebnisse	268
12.7.1. Eindimensionalität der Skalen	268
12.7.2. Exkurs: Raschanalyse	270
12.7.3. Zuverlässigkeit der Skalen	270
12.7.4. Eignung des Antwortformates	273
12.7.5. Messinvarianz	274
12.7.6. Dimensionsanalyse	277
12.7.7. Abgrenzbarkeit zu anderen Konstrukten	279
12.7.8. Abbildung theoretischer Erwartungen	281
12.8. Interpretation der Ergebnisse und Validitätsargumente	285
12.8.1. Eindimensionalität der Skalen	285
12.8.2. Zuverlässigkeit der Skalen	285
12.8.3. Eignung des Antwortformates	287
12.8.4. Messinvarianz	287

12.8.5. Dimensionsanalyse	288
12.8.6. Abgrenzbarkeit zu anderen Konstrukten	289
12.8.7. Abbildung theoretischer Erwartungen	289
12.9. Externe Validierung	290
12.10. Zusammenfassung	290

IV Schluss

13. Zusammenfassung, Diskussion und Ausblick	293
13.1. Zusammenfassung der Ergebnisse	293
13.1.1. Ergebnisse auf theoretischer Ebene	293
13.1.2. Ergebnisse auf empirischer Ebene in Bezug auf die Forschungs- fragen	296
13.1.3. Das wichtigste Ergebnis: Das Instrument	302
13.2. Diskussion	303
13.2.1. Zur Reflexion der Itementwicklung und -überarbeitung	303
13.2.2. Zum übergeordneten Validitätsargument	303
13.2.3. Zur Subjektivität und einem mittleren Spezifitätsniveau	305
13.3. Ausblick	307

V Anhang

A. Auswahl an Instrumenten zur Erfassung des Selbstkonzeptes	311
A.1. Instrumente zur Erfassung des Selbstkonzeptes von Schülerinnen und Schülern.	311
A.2. Instrumente zur Erfassung des Selbstkonzeptes von (angehenden) Lehre- rinnen und Lehrern.	314
B. Itemsammlung selten verwendeter Instrumente zur Erfassung von „Leh- rer-Selbstwirksamkeitserwartungen“	315
C. Exemplarische Darstellung der Bedeutung des Konstruktes ohne Berück- sichtigung der Erhebungsinstrumente	317
D. Itemsammlung selten verwendeter Instrumente zur Erfassung von „science teaching self-efficacy beliefs“	319
E. Vergleich der „Teacher Efficacy Scale“ und des „Science Teaching Effi- cacy Beliefs Instrument“	320

F. Physikdidaktische Themenfelder	321
G. Skalenversionen	324
G.1. Skalenversion 1 (V1): Pilotstudie I	324
G.1.1. Skala zum Experimentieren/Planung	324
G.1.2. Skala zum Experimentieren/Durchführung	324
G.1.3. Skala zum Elementarisieren/Planung	325
G.1.4. Skala zum Elementarisieren/Durchführung	325
G.1.5. Skala zum Umgang mit Schülervorstellungen/Planung	325
G.1.6. Skala zum Umgang mit Schülervorstellungen/Durchführung . .	326
G.2. Skalenversion 2 (V2): Pilotstudie II und III	327
G.2.1. Skala zum Experimentieren/Planung	327
G.2.2. Skala zum Experimentieren/Durchführung	327
G.2.3. Skala zum Elementarisieren/Planung	327
G.2.4. Skala zum Elementarisieren/Durchführung	328
G.2.5. Skala zum Umgang mit Schülervorstellungen/Planung	328
G.2.6. Skala zum Umgang mit Schülervorstellungen/Durchführung . .	329
G.2.7. Skala zum Umgang mit Aufgaben/Planung	329
G.2.8. Skala zum Umgang mit Aufgaben/Durchführung	330
G.3. Skalenversion 3 (V3): Hauptstudie	331
G.3.1. Skala zum Experimentieren/Planung	331
G.3.2. Skala zum Experimentieren/Durchführung	332
G.3.3. Skala zum Elementarisieren/Planung	333
G.3.4. Skala zum Elementarisieren/Durchführung	334
G.3.5. Skala zum Umgang mit Schülervorstellungen/Planung	335
G.3.6. Skala zum Umgang mit Schülervorstellungen/Durchführung . .	336
G.3.7. Skala zum Umgang mit Aufgaben/Planung	337
G.3.8. Skala zum Umgang mit Aufgaben/Durchführung	338
G.4. Skalenversion 4 (V4): Zur Verwendung empfohlen	339
G.4.1. Skala zum Experimentieren/Planung	339
G.4.2. Skala zum Experimentieren/Durchführung	340
G.4.3. Skala zum Elementarisieren/Planung	341
G.4.4. Skala zum Elementarisieren/Durchführung	342
G.4.5. Skala zum Umgang mit Schülervorstellungen/Planung	343
G.4.6. Skala zum Umgang mit Schülervorstellungen/Durchführung . .	344
G.4.7. Skala zum Umgang mit Aufgaben/Planung	345
G.4.8. Skala zum Umgang mit Aufgaben/Durchführung	346
H. Zum Interview	347
H.1. Skalen – Interviewversion	347

H.2.	Interviewleitfaden	348
H.2.1.	Phase I – Anmerkungen zur Einführung	348
H.2.2.	Phase II – Mögliche Impulse bez. des Vierschritts	348
H.2.3.	Phase III – Anmerkungen zur Gesamtskala	349
H.3.	Beispiel Interviewnotizen	351
H.4.	Beispielübersicht Itemauswertung (Interviews)	352
I.	Zur Expertenbefragung	355
I.1.	Expertenfragebogen (Physikdidaktik)	355
I.2.	Expertenfragebogen (Psychologie)	360
I.3.	Beispielübersicht Itemauswertung (Expertenbefragung)	365
I.4.	Beispielübersicht Skalenauswertung (Expertenbefragung)	366
J.	Fragebogen zur Akzeptanzbefragung bez. des Antwortformates	367
K.	Beispiel einer zusammenfassenden Übersicht zur Itemüberarbeitung nach Pilotstudien II & III	369
L.	Fragebogen – Hauptstudie	370
M.	Anreizsysteme und Aufrufe	376
M.1.	Flyer	376
M.2.	Foreneinträge	377
N.	Dokumentation Datenselektion	379
N.1.	Dokumentation für die Stichprobe der Lehrkräfte	379
N.2.	Dokumentation für die Stichprobe der Referendare	380
N.3.	Dokumentation für die Stichprobe der Studierenden	381
O.	Messinvarianz – zusätzliche Skalen	383
O.1.	Mehrgruppenvergleich: Personengruppe	383
O.2.	Mehrgruppenvergleich: Geschlecht	386
O.3.	Mehrgruppenvergleich: Studiengang und Schulform	387
O.4.	Mehrgruppenvergleich: Praxiserfahrung	388
P.	Mittelwertvergleiche	389
P.1.	Zusätzliche Skalen	389
P.2.	Neukonstruierte Skalen	391
Literatur		392
Skalendokumentation		420