

Die Überprüfung der transaktionalen Stresstheorie im Lehramtsreferendariat

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem
Fachbereich Psychologie
der Philipps-Universität Marburg
vorgelegt

von
Oliver Christ
aus Homberg/Ohm

Marburg/Lahn 2004

Vom Fachbereich Psychologie der Philipps-Universität Marburg als Dissertation
am 22.12.2004 angenommen.

Erstgutachter: Prof. Dr. Ulrich Wagner

Zweitgutachter: Prof. Dr. Peter Schmidt

Tag der mündlichen Prüfung am 20.01.2005.

Für meine lieben Eltern, Inge und Klaus Christ

Danksagung

Ich möchte mich an dieser Stelle bei wichtigen Personen in meinem Umfeld bedanken. Wie schon tausendfach in anderen Dissertationen angemerkt, kann das Vorhaben zu promovieren nur gelingen, wenn man sich der Unterstützung vieler Menschen sicher sein kann.

Dies sind an erster Stelle meine Eltern, Inge und Klaus Christ, ohne deren Unterstützung diese Arbeit nie möglich gewesen wäre und die immer für mich da sind. Ich hoffe, dass sich Ihr Vertrauen in mich mit dieser Arbeit ausgezahlt hat. Auch meiner Schwester Andrea möchte danken. Nicht nur für das sorgfältige Korrekturlesen, sondern auch für jegliche Art von Unterstützung, die sie mir hat zukommen lassen. Nicht unerwähnt bleiben darf Oma Lina! Ohne ihre Unterstützung hätte ich nicht über ausreichende „Hardware“ verfügt.

Besonderer Dank gilt Uli Wagner für die kompetente und angenehme Anleitung meiner Arbeit. Großen Dank auch für seine Toleranz gegenüber hoffnungsloser Unordnung!

Ein großes Dankeschön an Rolf van Dick, der mich gerade am Anfang sehr unterstützt hat und ganz wesentlich bei der Entwicklung der Ideen mit beteiligt war.

Peter Schmidt möchte ich für seine wertvollen Methoden-Tipps und seine menschliche Anleitung danken.

Jost Stellmacher und Andreas Homburg, meine geduldigen Kollegen, möchte ich dafür danken, dass sie meine Launen immer ausgehalten haben.

Schließlich möchte ich mich ganz besonders bei Julia bedanken. Es ist die schönste Erfahrung festzustellen, wie unwichtig eine Dissertation eigentlich ist! Vielen Dank dafür und für vieles mehr!

Inhaltsverzeichnis

Zusammenfassung	15
1 Einleitung	17
2 Stress: Definitive Ansätze	19
2.1 <i>Stress als Reizvariable</i>	19
2.2 <i>Stress als Reaktionsvariable</i>	20
2.3 <i>Stress als Beziehungskonzept</i>	21
3 Die transaktionale Stresstheorie	23
3.1 <i>Bewertungsprozesse</i>	25
3.2 <i>Stressbewältigung</i>	27
3.3 <i>Antezedentien und Ergebnisse des Stressprozesses</i>	30
3.4 <i>Wirkungszusammenhänge der Modellkomponenten</i>	32
3.5 <i>Weiterentwicklung der transaktionalen Stresstheorie</i>	33
4 Die empirische Überprüfung der transaktionalen Stresstheorie	35
4.1 <i>Kriterien für eine empirische Überprüfung der transaktionalen Stresstheorie</i>	35
4.2 <i>Die transaktionale Stresstheorie in ihrer empirischen Bewährung</i>	36
4.2.1 <i>Abbildung intraindividuelle Veränderungen</i>	38
4.2.2 <i>Korrelative Untersuchungen</i>	39
4.2.3 <i>„Examens-Studien“</i>	43
4.2.4 <i>Experimentelle Untersuchungen</i>	46
4.2.5 <i>Pfad- und Strukturgleichungsanalysen</i>	48
4.2.6 <i>Kombination intra- und interindividueller Analyseverfahren</i>	53
4.3 <i>Stand der empirischen Überprüfung der transaktionalen Stresstheorie: Ein Fazit</i>	54
5 Kontext und Fragestellung der Arbeit	57
5.1 <i>Kontext der Arbeit: Das Lehramtsreferendariat</i>	57
5.1.1 <i>Struktur und inhaltlicher Aufbau des Lehramtsreferendariats</i>	58
5.1.2 <i>Belastung und Beanspruchung von ReferendarInnen</i>	59
5.1.3 <i>Belastung und Beanspruchung von LehrerInnen</i>	61

5.1.4	Anwendung des Stressmodells in der Lehrerstress-Forschung.....	62
5.1.5	Das Lehramtsreferendariat als Stressor.....	63
5.2	<i>Fragestellung der Arbeit</i>	64
6	Studie 1	71
6.1	<i>Methode Studie 1</i>	71
6.1.1	Vorgehen.....	71
6.1.2	Stichprobe.....	72
6.1.3	Messinstrument.....	74
6.2	<i>Ergebnisse Studie 1</i>	82
6.2.1	Allgemeines Vorgehen.....	82
6.2.2	Datenscreening.....	84
6.2.3	Datenanalyse - Vorschritte.....	84
6.2.4	Skalenüberprüfung.....	90
6.2.5	Überprüfung des Messmodells.....	101
6.2.6	Modellüberprüfung.....	113
6.3	<i>Diskussion Studie 1</i>	125
7	Studie 2	127
7.1	<i>Methode Studie 2</i>	127
7.1.1	Vorgehen.....	127
7.1.2	Stichprobe.....	129
7.1.3	Messinstrument.....	131
7.2	<i>Ergebnisse Studie 2</i>	132
7.2.1	Allgemeines Vorgehen bei der Datenanalyse.....	132
7.2.2	Panel-Mortalität.....	138
7.2.3	Datenaufbereitung.....	141
7.2.4	Überprüfung des Messmodells.....	145
7.2.5	Überprüfung des Stressmodells im Querschnitt.....	154
7.2.6	Überprüfung des Stressmodells im Längsschnitt.....	157
7.2.7	Zusammenfassendes Fazit der Ergebnisse aus Studie 2.....	180
7.2.8	Zusatzanalysen.....	182
7.3	<i>Diskussion Studie 2</i>	184

8	Gesamtdiskussion	187
8.1	<i>Die Bewährung der transaktionalen Stresstheorie</i>	187
8.1.1	Zusammenfassung der Ergebnisse	187
8.1.2	Diskussion der Befunde	189
8.2	<i>Abschließendes Fazit und Ausblick.....</i>	194
8.3	<i>Implikationen für eine theoriegeleitete Stressprävention.....</i>	196
9	Literatur	201
10	Anhang	I

Tabellen- und Abbildungsverzeichnis

<i>Tabelle 1.</i>	Ausbildungsabschnitte und deren Dauer im Lehramtsreferendariat	58
<i>Tabelle 2.</i>	Verteilung der Befragten auf die verschiedenen Studienseminarformen im Vergleich zu Hessen	73
<i>Tabelle 3.</i>	Geschlechtsverteilung der Befragten für die einzelnen Studienseminarformen im Vergleich zu Hessen	73
<i>Tabelle 4.</i>	Konstrukte und deren Operationalisierung	75
<i>Tabelle 5.</i>	Lehrertypische und ausbildungsspezifische Belastungsfaktoren (Reihenfolge wie im Fragebogen)	77
<i>Tabelle 6.</i>	Messinstrument zur Erfassung sozialer Unterstützung	79
<i>Tabelle 7.</i>	Geschlechtsverteilung und Aufteilung auf die unterschiedlichen Ausbildungsphasen und Studienseminare (in Klammern in Prozent) in der Kalibrierungs- und Validierungsstichprobe	85
<i>Tabelle 8.</i>	Vergleich der beiden Teilstichproben in Hinblick auf weitere soziodemographische Variablen	86
<i>Tabelle 9.</i>	Mittelwert und in Klammern Spannbreite der univariaten Schiefe und Kurtosis in den zentralen Modellvariablen für die Kalibrierungs- und Validierungsstichprobe	89
<i>Tabelle 10.</i>	Ladungsmuster (nach Rotation) der Items der Belastungsskala auf den vier extrahierten Faktoren (Faktorladungen von $a < .20$ sind nicht aufgeführt; Items mit fett markierter Ladung werden ausgewählt)	92
<i>Tabelle 11.</i>	Interkorrelation der vier Belastungsfaktoren	93
<i>Tabelle 12.</i>	Faktorladungen und Interkorrelationen des Messmodells (Kalibrierungsstichprobe)	102
<i>Tabelle 13.</i>	Faktorladungen und Interkorrelationen des Messmodells (Validierungsstichprobe)	104
<i>Tabelle 14.</i>	Fit-Indizes der Modelle der Multigruppenanalyse zur Kreuzvalidierung des Messmodells	107
<i>Tabelle 15.</i>	Mittelwertsschätzung, Standardmessfehler der Strukturvariablen für die Validierungsstichprobe	108
<i>Tabelle 16.</i>	Mittelwerte in den Strukturvariablen für beide Stichproben	109

<i>Tabelle 17.</i>	Mittelwerte und Standardabweichungen für die Belastungsskalen (Kalibrierungsstichprobe).....	110
<i>Tabelle 18.</i>	Einfache Korrelationen zwischen den Belastungsskalen und den Modellvariablen.....	110
<i>Tabelle 19.</i>	Mittelwerte und Standardabweichungen für die Ausbildungsphasen der Kalibrierungsstichprobe.....	112
<i>Tabelle 20.</i>	Fit-Indizes der Modelle der Multigruppenanalyse	120
<i>Tabelle 21.</i>	Fit-Werte des modifizierten Stressmodells für die weiblichen und männlichen Befragten sowie getrennt für die Ausbildungsphasen (Kalibrierungsstichprobe)	122
<i>Tabelle 22.</i>	Fit-Werte der Modelle der Multigruppenanalyse zum Vergleich der weiblichen und männlichen Befragten der Kalibrierungsstichprobe	123
<i>Tabelle 23.</i>	Fit-Werte der Modelle der Multigruppenanalyse zum Vergleich der Ausbildungsphasen der Kalibrierungsstichprobe	124
<i>Tabelle 24.</i>	Zeitlicher Ablauf von Studie 2 und Ausbildungsstand der befragten ReferendarInnen zu den jeweiligen Messzeitpunkten	128
<i>Tabelle 25.</i>	Anzahl der Befragten und Rücklaufquoten für t2, t3 und t4 (in Klammern Rück-laufquote in %)......	129
<i>Tabelle 26.</i>	Beschreibung der drei Datensätze L1, L2, L3 in zentralen soziodemographischen Variablen zum ersten Messzeitpunkt	130
<i>Tabelle 27.</i>	Über die Fragestellung hinausgehende Skalen zu den vier Messzeitpunkten	131
<i>Tabelle 28.</i>	Verteilung der Befragten auf Geschlecht, Studienseminar, Familienstand und Kinder (in Klammern in Prozent) getrennt nach Dropouts und Continuers...	140
<i>Tabelle 29.</i>	Fit-Indizes der Modelle der Multigruppenanalyse zur Überprüfung der Übereinstimmung in den Mittelwerten, Varianzen und Kovarianzen zwischen den Dropouts und den Continuern.....	141
<i>Tabelle 30.</i>	Anteil an Dropouts (in Klammern in Prozent) zu allen Messzeitpunkten für die Datensätze L2 und L3	142
<i>Tabelle 31.</i>	Mittlere Schiefe und Kurtosis (in Klammern Standardabweichung) in den zentralen Modellvariablen sowie Mardia-Koeffizient zu allen Messzeitpunkten getrennt für die Datensätze L1, L2 und L3.....	144

<i>Tabelle 32.</i>	Ergebnisse der CFAs zur Überprüfung der Messmodelle zu allen Messzeitpunkten getrennt nach Datensätzen.....	147
<i>Tabelle 33.</i>	Ergebnisse für Modell 1 (Test auf konfigurale Invarianz) der Multigruppenanalyse 1 für alle Datensätze	149
<i>Tabelle 34.</i>	Ergebnisse für Modell 1 (Test auf konfigurale Invarianz) der Multigruppenanalyse 2 für alle Datensätze	149
<i>Tabelle 35.</i>	Ergebnisse der Berechnung der Messmodelle mit der Bootstrap-Methode für alle Messzeitpunkte und alle Datensätze	151
<i>Tabelle 36.</i>	Ergebnisse der Multigruppenanalyse zur Überprüfung von Verzerrungseffekten durch unterschiedliche Typen von Continuern.....	153
<i>Tabelle 37.</i>	Ergebnisse der Überprüfung des Stressmodells zu den Messzeitpunkten 2, 3 und 4 unter Verwendung von SEM für alle drei Datensätze	154
<i>Tabelle 38.</i>	Ergebnisse der Multigruppenanalyse zum Vergleich der drei Längsschnittdatensätze (L1, L2, L3) zum Messzeitpunkt 2 und der Kalibrierungsstichprobe aus Studie 1	155
<i>Tabelle 39.</i>	Ergebnisse für Modell 1 der Multigruppenanalyse zum Test auf Stabilität des (querschnittlichen) Stressmodells für alle Datensätze.....	157
<i>Tabelle 40.</i>	Mittelwerte und Standardabweichungen für die Modellvariablen zu allen Messzeitpunkten (Datensatz L2).....	160
<i>Tabelle 41.</i>	Fit-Indizes für das restringierte Modell zur Überprüfung der Stabilität der Mittelwerte in allen Modellvariablen über die Zeit.....	161
<i>Tabelle 42.</i>	Fit-Indizes für das ARCL-Modell jeweils getrennt für die Modellvariablen (Modelle ohne Restriktionen).....	162
<i>Tabelle 43.</i>	Fit-Indizes für das einfaktorielle LGC-Modell jeweils getrennt für die Modellvariablen.....	163
<i>Tabelle 44.</i>	Fit-Indizes für das zweifaktorielle LGC-Modell jeweils getrennt für die Modellvariablen.....	164
<i>Tabelle 45.</i>	Fit-Indizes für das ALT-Modell jeweils getrennt für die Modellvariablen....	166
<i>Tabelle 46.</i>	Vergleich von ARCL-, ein- und zweifaktoriellem LGC-Modell und ALT-Modell für alle Modellvariablen.....	167

<i>Tabelle 47.</i>	Fit-Indizes für die zu jedem Analyseschritt optimierten ALT-Modelle (Datensatz L2)	171
<i>Tabelle 48.</i>	Intercepts und Varianzen der latenten Intercept- und Slope-Faktoren sowie der Skala Neurotizismus des multivariaten ALT-Modells (Datensatz L2)	173
<i>Tabelle 49.</i>	Ergebnisse der Multigruppenanalyse zum Vergleich des multivariaten ALT-Modells in allen drei Datensätzen	175
<i>Tabelle 50.</i>	Ergebnisse der Berechnung des multivariaten ALT-Modells mit der Bootstrap-Methode für alle Datensätze (unstandardisierte Parameter).....	176
<i>Tabelle 51.</i>	Einfluss von Geschlecht auf die latenten Modellvariablen des ALT-Modells	178
<i>Abbildung 1.</i>	Schematische Darstellung der transaktionalen Stresstheorie (in Anlehnung an Schwarzer, 2000, S. 16).....	24
<i>Abbildung 2.</i>	Schematische Darstellung der Untersuchungshypothesen	67
<i>Abbildung 3.</i>	Das Grundmodell in seiner empirischen Umsetzung	114
<i>Abbildung 4.</i>	Das modifizierte Stressmodell mit allen Strukturgleichungen, Korrelationen und der aufgeklärten Varianz der endogenen Faktoren für die Kalibrierungsstichprobe.....	116
<i>Abbildung 5.</i>	Das modifizierte Stressmodell mit allen Strukturgleichungen, Korrelationen und der aufgeklärten Varianz der endogenen Faktoren für die Validierungsstichprobe.....	118
<i>Abbildung 6.</i>	Ein univariates LGC-Modell für eine Variable, die zu vier Messzeitpunkten erhoben wurde (y1 bis y4)	135
<i>Abbildung 7.</i>	Ein univariates ALT-Modell für eine Variable, die zu vier Messzeitpunkten erhoben wurde (y1 bis y4)	137
<i>Abbildung 8.</i>	Beispiel für ein ARCL-Modell (Simplex-Modell) in Schritt 2 mit vier Wiederholungsmaßen	158
<i>Abbildung 9.</i>	Schematische Darstellung der strukturellen Beziehungen zwischen den Modellvariablen für das multivariate ALT-Modell (standardisierte Parameter)	172
<i>Abbildung 10.</i>	Zusammenfassende Darstellung der zentralen Untersuchungsergebnisse	188

Zusammenfassung

Ziel der Arbeit ist die Überprüfung der transaktionalen Stresstheorie von R.S. Lazarus (1966; Lazarus & Folkman, 1984). Die Theorie definiert Stress als eine Transaktion zwischen einer Person und der Umwelt. Stress entsteht dann, wenn eine Person eine Situation als Gefahr für ihr Wohlbefinden bewertet und gleichzeitig die Situation Anforderungen an die Person stellt, die ihre Ressourcen beanspruchen oder übersteigen. Kurz- und langfristige Folgen dieser Transaktion sind abhängig von den Bewältigungsstrategien, die eine Person auf Basis der Bewertung der Situation auswählt. Neben den Annahmen zur Struktur der Variablen wird die Person-Umwelt-Relation dynamisch konzipiert. Das Bewältigungsverhalten wirkt auf die Bewertungsprozesse zurück, eine Person beeinflusst aktiv die Person-Umwelt-Relation. Sowohl die Struktur als auch der Prozess müssen bei einer Überprüfung der Theorie berücksichtigt werden. Es finden sich eine Reihe von Untersuchungen, die direkt oder indirekt die transaktionale Stresstheorie überprüft haben. Gemeinsam ist allen Studien ihre unzureichende methodische Herangehensweise, die eine adäquate Testung der Theorie nicht zulässt. Die Hauptmängel bestehen darin, dass entweder nur interindividuelle (Struktur) oder nur intraindividuelle Varianz (Prozess) bei der Auswertung berücksichtigt wurden, oder, wenn beide Varianzquellen gleichzeitig in die Analyse eingingen (Gunthert, Cohen & Armeli, 1999), die strukturelle Beziehung der Modellkomponenten nicht überprüft wurde.

Die Überprüfung der transaktionalen Stresstheorie erfolgt in der vorliegenden Arbeit im Kontext des Lehramtsreferendariats, es wurden eine Quer- (Studie 1) und eine Längsschnittbefragung (Studie 2) realisiert. Mittels eines standardisierten Fragebogens wurden in beiden Studien Angaben zu Neurotizismus, zum Ausmaß an sozialer Unterstützung, zur Ereignis- und Ressourceneinschätzung, zum Bewältigungsverhalten, zum psychischen Wohlbefinden und zum Ausmaß von körperlichen Beschwerden erhoben. Strukturgleichungsanalysen können in Studie 1 (N = 645 ReferendarInnen) nur in Teilen die strukturellen Annahmen der transaktionalen Stresstheorie bestätigen. Neurotizismus und das Ausmaß an sozialer Unterstützung wirken auf die Ereignis- und Ressourceneinschätzung, die wiederum das Bewältigungsverhalten beeinflussen. Das Bewältigungsverhalten hat aber nur einen Effekt auf das psychische Wohlbefinden, körperliche Beschwerden stehen in keinem Zusammenhang mit der Bewältigung. Darüber hinaus beeinflusst Neurotizismus entgegen den Erwartungen direkt das psychische Wohlbefinden und das Ausmaß an Beschwerden. Die prozessualen Annahmen können mit einer Querschnittsuntersuchung nicht überprüft werden, Längsschnittuntersuchungen ermöglichen dagegen die Analyse des Prozesses. In Studie 2 (N = 398 ReferendarInnen)

wurden die Modellkomponenten zu vier Messzeitpunkten, die über einen Zeitraum von 13 Monaten verteilt waren, erhoben. Zur Auswertung der Längsschnittdaten wurde auf das sogenannte Autoregressive-Latent-Trajectory (ALT)-Modell (Bollen & Curran, 2004; Curran & Bollen, 2001) zurückgegriffen. Dieses Verfahren erlaubt es, sowohl interindividuelle Unterschiede als auch intraindividuelle Veränderungen in den Modellkomponenten zu analysieren. Darüber hinaus können die Beziehungen der Komponenten über die Zeit berücksichtigt werden. Auch in Studie 2 können die Annahmen der transaktionalen Stresstheorie nur in Teilen bestätigt werden. Nur Unterschiede und Veränderungen im psychischen Wohlbefinden können durch das Bewältigungsverhalten erklärt werden, das Ausmaß an körperlichen Beschwerden ist unabhängig von der Bewältigung. Ein dynamisches Wechselspiel zwischen dem Bewältigungsverhalten und den Bewertungsprozessen lässt sich nicht finden. Die Ereignis- und Ressourceneinschätzung bleiben über den Erhebungszeitraum stabil, Veränderungen im Bewältigungsverhalten klären keine Varianz in den Ergebnisvariablen, psychisches Wohlbefinden und Ausmaß an körperlichen Beschwerden, auf. Es finden sich keine rekursiven Beziehungen zwischen dem Bewältigungsverhalten und der Ereignis- und Ressourceneinschätzung.

Die Ergebnisse beider Studien zeigen nur eine geringe Bestätigung der Annahmen der transaktionalen Stresstheorie. Kognitive Bewertungsprozesse und das Bewältigungsverhalten haben nur auf das psychische Wohlbefinden einen Einfluss. Die zentralen Prozessvariablen der Stresstheorie, die Ereignis- und Ressourceneinschätzung und das Bewältigungsverhalten, verändern sich nicht über die Zeit, sondern spiegeln stabile, interindividuelle Unterschiede wieder. Möglicherweise lassen sich dynamische Prozesse nur für kleinere Zeitfenster finden. Zukünftige Untersuchungen sollten kürzere und längere Zeitabschnitte verknüpfen, um dynamische und stabile Muster im Umgang mit stressigen Situationen trennen zu können. Das in Studie 2 gewählte Analyseverfahren kann dabei helfen, solche komplexen Beziehungsmuster zu finden.

1 Einleitung

Stress ist in unserem Alltag ein häufig gebrauchter Begriff. Von Stress wird gesprochen, wann immer sich Personen überlastet und überfordert fühlen (Nitsch, 1981a). Es gibt kaum Lebensbereiche, die nicht im Zusammenhang mit dem Begriff Stress stehen. Ob an der Arbeit, in der Familie oder in der Freizeit: Der Begriff Stress wird immer dann verwendet, wenn das körperliche und psychische Wohlbefinden beeinträchtigt ist. Teilweise wird bereits von einer Stressgesellschaft gesprochen (Fritzsche, 1998), wobei ein Zusammenhang zwischen individuellem Stress und der Modernisierung bzw. den schnell wandelnden Gesellschaften gesehen wird.

Klar ist, dass Stress als Phänomen eine hohe gesellschaftliche Relevanz besitzt und für viele Personen ein wichtiges, meist negatives Empfinden darstellt. Auch in der Wissenschaft ist Stress ein viel untersuchtes Phänomen. So kommen Hobfoll, Schwarzer und Chon (1998) zu dem Ergebnis, dass Stress und dessen Bewältigung das wahrscheinlich am meisten untersuchte Phänomen in der heutigen wissenschaftlichen Psychologie ist.

Stress erlebt jeder Mensch, dennoch gibt es große interindividuelle Unterschiede. Es gibt zwar theoretische Ansätze, die versuchen, bestimmte Situationen zu finden, die für alle Personen ein mehr oder minder stressiges Ereignis darstellen. Doch stellen objektiv gleiche Situationen für die eine Person ein extremes Ereignis dar, welches mit bedeutsamen Konsequenzen für das Wohlbefinden einhergeht. Für eine andere Person dagegen ist es ein Ereignis, über das es sich kaum lohnt nachzudenken. Das Stressausmaß, welches eine Situation bei einem Menschen auslöst, ist damit nicht alleine von Merkmalen der Situation abhängig, sondern auch davon, wie eine Person die Situation bewertet und mit der Situation umgeht. Genau dies sind die zentralen Bestandteile der transaktionalen Stresstheorie von Richard S. Lazarus (Lazarus, 1991a; Lazarus & Folkman, 1984). Die Theorie fokussiert solche Bewertungs- und Bewältigungsprozesse, um einerseits Unterschiede zwischen Personen, andererseits aber auch den Stressprozess, den eine Person bei der Konfrontation mit einer stressigen Situation durchläuft, zu erklären.

Die transaktionale Stresstheorie wird häufig im Zusammenhang mit dem Phänomen Stress zitiert, sie bildet die theoretische Grundlage für eine Vielzahl von wissenschaftlichen Untersuchungen und auch Stressbewältigungstrainings. Trotz ihrer Beliebtheit und augenscheinlichen Plausibilität ist bislang die Prüfung der Theorie Gegenstand von nur wenigen empirischen Arbeiten gewesen. Die vorliegende Arbeit möchte genau an dieser Lücke ansetzen und im Rahmen einer umgrenzten Situation, dem Lehramtsreferendariat, die Theorie und

ihre Vorhersagen überprüfen. Die Arbeit gliedert sich in zwei große Teile, einem theoretischen Teil (Kapitel 2 bis 5) und einem empirischen Teil (Kapitel 6 und 7). Im Theorieteil der Arbeit wird zunächst ein Überblick über wissenschaftliche Definitionen von Stress gegeben und schließlich das der Arbeit zu Grunde liegende Begriffsverständnis von Stress vorgestellt (Kap. 2). Im Anschluss daran wird die transaktionale Stresstheorie von Richard S. Lazarus vorgestellt, deren empirische Prüfung Gegenstand der vorliegenden Arbeit ist (Kap. 3). Die Arbeit stellt nicht die erste empirische Prüfung der Theorie dar. Eine Reihe von Untersuchungen wurde bereits durchgeführt mit dem zum Teil expliziten Ziel, die strukturellen und prozessualen Annahmen der Theorie zu überprüfen. Diese Untersuchungen werden in Kapitel 4 einer kritischen Bewertung unterzogen. Der Theorieteil endet mit der Beschreibung des Untersuchungskontexts und der genauen Explikation der Fragestellung (Kap. 5). In Rahmen zweier Studien wird die Überprüfung der Theorie vorgenommen, einer Querschnittuntersuchung (Studie 1; Kap. 6) und einer Längsschnittuntersuchung (Studie 2, Kap. 7). Die Arbeit wird schließlich mit einer Gesamtdiskussion abgeschlossen (Kap. 8). Neben der Bewertung der Ergebnisse der beiden Studien wird ein Ausblick für weitere Forschung gegeben und auf Implikationen der Arbeit für die Prävention von Stress, insbesondere im Lehramtsreferendariat, eingegangen.

2 Stress: Definitiorische Ansätze

Während Stress umgangssprachlich relativ einheitlich verwendet wird, gibt es in der Wissenschaft eine große Heterogenität hinsichtlich der Begriffsbestimmung. Nach Cohen, Kessler und Gordon (1995; s.a. Kaluza & Vögele, 1999; Laux, 1983; Lazarus, 1991a; Lazarus & Folkman, 1984; Nitsch, 1981a) kann man zumindest zwischen drei verschiedenen Definitionsansätzen unterscheiden, denen jeweils unterschiedliche Modellvorstellungen zu Grunde liegen: (1) Der reizorientierte Ansatz, der aus den Umweltwissenschaften stammt, (2) der reaktionsorientierte Ansatz, der einer biologischen Tradition entstammt und schließlich (3) der transaktionale bzw. zustandsorientierte Ansatz, der seinen Ursprung in der psychologischen Stressforschung findet. Im Folgenden werden die entsprechenden Ansätze kurz skizziert sowie die Hauptkritikpunkte angeführt. Der der Arbeit zugrunde liegende Ansatz wird ausführlicher vorgestellt.

2.1 *Stress als Reizvariable*

Im reizorientierten Definitionsansatz wird Stress als ein situatives oder Reizphänomen angesehen, wobei Stress dann im Sinne einer unabhängigen Variablen über bestimmte Reiz-, Situations-, Bedingungs-, Ereignis- oder Umweltmerkmale operationalisiert wird. Ein prominentes Beispiel für die situationistische Sichtweise von Stress ist die klassische Life-Event-Forschung (Dohrenwend & Dohrenwend, 1974), wenn auch neuere Ansätze die rein situationsbezogene Sichtweise um eine subjektive Bewertungsdimension erweitert haben (Filipp, 1995). Hier wird Stress durch die Anzahl an bestimmten Lebensereignissen (z.B. Tod einer nahe stehenden Person, Unfälle) quantifiziert (für einen klassischen Ansatz s. Holmes & Rahe, 1967) und in Verbindung mit Krankheitsentwicklungen gesetzt. Der rein situationsspezifische Ansatz birgt eine Reihe von Problemen und Schwierigkeiten, die sich nach Nitsch (1981a, S. 42; s.a. Nitsch, 1981b) wie folgt zusammenfassen lassen: (a) Qualitativ unterschiedliche Reize oder Ereignisse können nicht auf demselben Maßstab abgebildet werden; (b) Die Bedeutung unterschiedlicher Reizkombinationen kann nicht abgebildet werden, Wechselwirkungen zwischen verschiedenen Reizbedingungen werden nicht berücksichtigt; (c) Die subjektive Wahrnehmung und Bewertung von Reizen findet keine Berücksichtigung in den klassischen reizorientierten Ansätzen; (d) Die Kenntnis der Reizbedingungen ermöglicht keine Vorhersage bestimmter psychologischer und physiologischer Reaktionen, da auch in solchen Fällen, in denen Reizbedingungen allgemein als stressend empfunden werden, sich interindividuelle Reaktionsunterschiede in Abhängigkeit von den jeweiligen Stressbewälti-

gungsfähigkeiten und –mustern zeigen; (e) Eine rein reizorientierte Stressdefinition führt dazu, dass unterschiedlichste Reaktionen zu Stressreaktionen werden (z.B. Ärger, Belustigung, Leistungssteigerung und Leistungsabfall). Extreme physiologische und psychische Reaktionen werden somit zu Stressfolgen.

Die aufgelisteten Schwierigkeiten einer rein reizorientierten Stressdefinition machen deutlich, dass es wenig sinnvoll erscheint, Stress alleine über Reizbedingungen zu operationalisieren. Dennoch spielen situative Bedingungen eine wichtige Rolle im Stressprozess. In modernen Stressansätzen werden aus diesem Grund Situationsfaktoren als (potenzielle) *Stressoren* bezeichnet, womit gemeint ist, dass bestimmte Situationen bzw. Reizbedingungen durch unterschiedliche Wahrscheinlichkeiten gekennzeichnet sind, mit denen sie Funktionsstörungen auslösen können.

2.2 *Stress als Reaktionsvariable*

In reaktionsorientierten Definitionen wird Stress als abhängige Variable gesehen. In diesen Ansätzen wird angenommen, dass Stress durch bestimmte Ereignisse ausgelöst wird. Die Stressreaktion ist ablesbar durch umschriebene physiologische, psychische oder verhaltensmäßige Störungs- und Anpassungsreaktionen. Einen berühmten und einflussreichen reaktionsorientierten Ansatz stellt die Stresstheorie von Selye (1936, 1976a, 1976b) dar. Selye definiert Stress als komplexes, aber einheitliches Reaktionsmuster, welches unspezifisch auf jede Anforderung erfolgt. Stressoren (also bestimmte Reizbedingungen, s.o.) lösen nach Selye eine physiologische Stressreaktion aus, die einen universellen Abwehrmechanismus darstellt, der die biologische Integrität des Organismus schützt. Diese von Selye als Allgemeines Adaptationssyndrom bezeichnete Reaktion soll bzw. kann drei Phasen durchlaufen. Zunächst reagiert der Organismus auf eine Stressor-Exposition mit der so genannten Alarmreaktion, die durch eine erhöhte Ausschüttung von Katecholaminen und Corticosteroiden gekennzeichnet ist (Hypothalamus-Nebennierenmark-Achse, Hypothalamus-Hypophysen-Nebennierenrindensystem), um den Organismus für eine Reaktion auf den Stressor vorzubereiten. In der daran anschließenden Widerstandsphase zeigt der Organismus der Alarmphase gegensätzliche Reaktionen, um die Gleichgewichtslage wiederherzustellen. Dauert die stressauslösende Situation an, folgt die Erschöpfungsphase, die durch den Zusammenbruch des Widerstandes bis hin zum Tod gekennzeichnet ist.

Der Ansatz von Selye wie auch reaktionsorientierte Ansätze generell weisen eine Reihe von Problemen auf (Kaluza & Vögele, 1999; Nitsch, 1981a): (a) Die Annahme von Selye, dass das physiologische Reaktionsmuster unspezifisch ist, hat nicht bestätigt werden können

(Mason, 1971, 1975a, 1975b). Somit ist die Stressreaktion auch durch reizspezifische Effekte und überdauernde physiologische und psychische Merkmale der betreffenden Person bestimmt. Selye hat in Reaktion auf diese Befunde sein ursprüngliches Unspezifitäts-Konzept modifiziert (Selye, 1991); (b) Ein weiteres Problem ist die Mehrdeutigkeit der Unspezifitäts-Annahme. Welche Situation die entsprechende Stressreaktion ausgelöst hat, lässt sich nur ex post feststellen, Vorhersagen über die Wirkung von Reizkonstellationen, wissenschaftliche Prognosen und begründete Prävention sind nicht möglich.

Zwar ist es wichtig, bestimmte Reaktionen als (potenzielle) Stressindikatoren zu identifizieren, allerdings ist es kritisch, alleine aus einem bestimmten Reaktionsmuster auf Stress zu schließen, ohne die auslösenden Reize und vermittelnden Prozesse zu berücksichtigen (Nitsch, 1981a).

2.3 *Stress als Beziehungskonzept*

In relationalen Stresskonzeptionen (Chalmers, 1981; Cox, 1978; Cox & Mackay, 1981; Hobfoll, 1988, 1998; Lazarus, 1966; Lazarus & Folkman, 1984; Lazarus & Launier, 1978; McGrath, 1970; Schwarzer, 1986; Thomae, 1987) wird Stress verstanden als Ergebnis komplexer, subjektiver Einschätzungsprozesse bezüglich der eigenen Ziele, Handlungsalternativen und Situationsbedingungen. Stress entsteht dann, wenn individuelle Handlungsvoraussetzungen und situative Handlungsbedingungen instrumentell (Diskrepanz zwischen Fähigkeiten und Anforderungen) oder motivational (Diskrepanz zwischen Bedürfnissen und Befriedigungsmöglichkeiten) im Ungleichgewicht stehen (Nitsch, 1981a; S. 47).

Stress kann aus einer relationalen Perspektive mit Lazarus und Folkman (1984, S. 19) definiert werden als „...a particular relationship between the person and the environment that is appraised by the person as taxing or exceeding his or her resources and endangering his or her well-being.“

Gemein ist allen relationalen Stresskonzeptionen die Analyse einer Person-Umwelt-Beziehung, die durch psychologische Variablen vermittelt wird. Die stressrelevante Person-Umwelt-Beziehung wird teilweise als statisch gesehen im Sinne einer aktuellen Konfrontation von Person und Situation. Andere Ansätze berücksichtigen die gegenseitige Beeinflussung von Situation und Person (Jerusalem, 1990; s. ausführlich zur Unterscheidung von Interaktion und Transaktion Lazarus & Launier, 1978). Diese transaktionale Sichtweise von Stress liegt auch der prominenten und einflussreichen Stresstheorie von Lazarus und Mitarbeitern zu Grunde (Lazarus, 1966, 1991a; Lazarus & Folkman, 1984, 1987; Lazarus & Launier, 1978). Die transaktionale Stresstheorie soll in der vorliegenden Arbeit im Kontext einer spezifischen

Situation, dem Lehramtsreferendariat, einer empirischen Überprüfung unterzogen werden und wird im folgenden Kapitel ausführlicher vorgestellt.

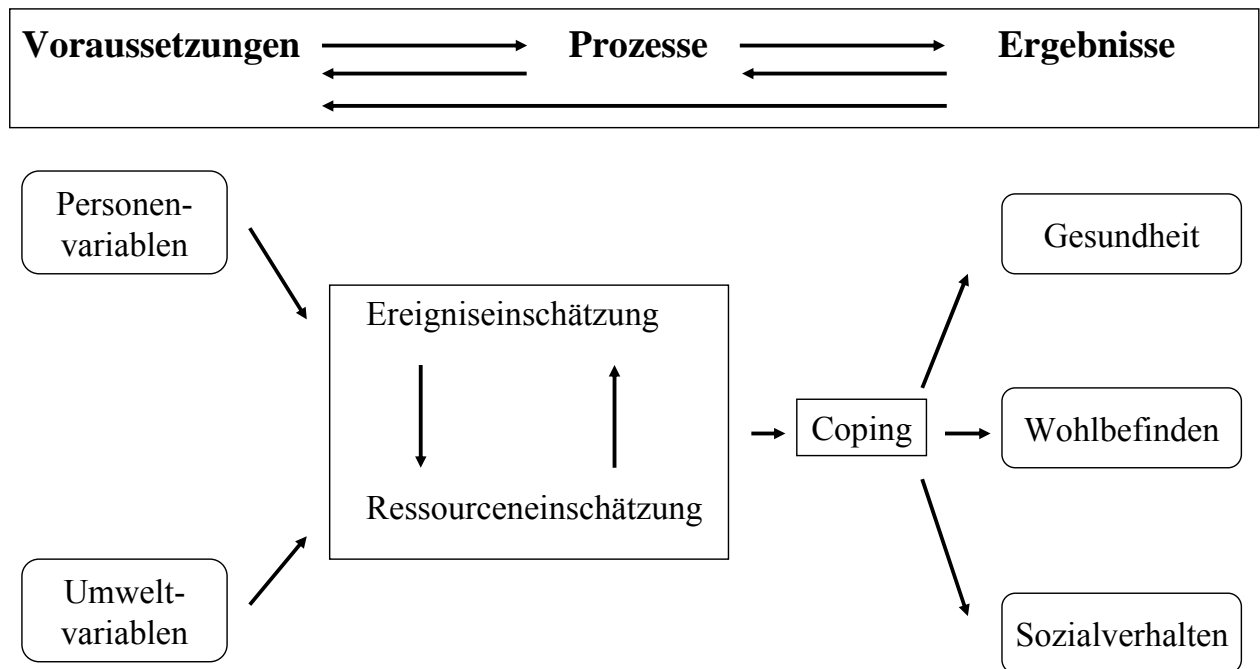
3 Die transaktionale Stresstheorie

Basierend auf einem umfangreichen Forschungsprogramm in den 50er und 60er Jahren des 20. Jahrhunderts und der darauf basierenden Beobachtung, dass es interindividuelle Unterschiede in den Stressreaktionen zwischen Personen gibt (zum Überblick s. Lazarus, 1968; Lazarus, Averill & Opton, 1970), formulierte Lazarus erstmals 1966 eine transaktionale Sichtweise von Stress, deren zentrale Bestandteile kognitive Bewertungen und Bewältigungsprozesse waren. Im Laufe der Zeit wurde diese Theorie von Lazarus und seinen Mitarbeitern weiter ausdifferenziert (Lazarus & Launier, 1978; Lazarus & Folkman, 1984, 1987). Schließlich fokussiert der späte Lazarus mehr auf Emotionen im Allgemeinen und baut seine Stresstheorie zu einem kognitiv-motivational-emotiven System aus (Lazarus, 1991a, 1991b, 1991c, 1993a, 1999; Smith & Lazarus, 1990, 1993). Lazarus sieht nunmehr (psychologischen) Stress nicht als eigenständiges psychologisches Phänomen an, sondern als einen Ausschnitt aus einer übergeordneten Klasse psychologischer Phänomene, nämlich den Emotionen. Seiner Ansicht nach sollte die Stress- und Emotionsforschung, die bis dahin mehr oder weniger nebeneinander existierten, zu einem Forschungszweig zusammengeführt werden. Lazarus' Interesse fokussiert nun auf spezifische Emotionen wie Eifersucht, Trauer, Ärger, Freude, etc. und deren spezifische Person-Umwelt-Beziehungen. Spezifische Emotionen sollen Aufschluss darüber liefern, welcher Art von Belastung eine Person ausgesetzt ist. Stress dagegen informiert nur allgemein darüber, dass die betroffene Person stark belastet ist. Somit stellt die Untersuchung von Emotionen eine feinere Analyse von Person-Umwelt-Relationen dar. Die transaktionale Sichtweise bleibt aber Kernstück dieser neuen Formulierungen. Die kognitiv-motivational-emotive Sichtweise von Emotionen ist keine grundsätzliche Neuformulierung der transaktionalen Stresstheorie, sondern eine Erweiterung auf Emotionen im Allgemeinen.

Ziel der vorliegenden Arbeit ist die Überprüfung der ursprünglichen transaktionalen Stresstheorie, die folgend ausführlich beschrieben wird. Der Gegenstandsbereich beschränkt sich auf Stress im Allgemeinen, spezifische Emotionen stehen nicht im Fokus der Arbeit.

In Abbildung 1 ist die transaktionale Theorie schematisch dargestellt (vgl. Lazarus & Folkman, 1987).

Abbildung 1. Schematische Darstellung der transaktionalen Stresstheorie (in Anlehnung an Schwarzer, 2000, S. 16)



Im Zentrum des Modells stehen die kognitiven Bewertungsprozesse (Ereignis- und Ressourceneinschätzung; „cognitive appraisal“; vgl. Arnold, 1960) und die Stressbewältigung („coping“). Kognitive Bewertungen werden dabei verstanden als kontinuierliche Überprüfung der Umweltgegebenheiten daraufhin, ob sie in irgendeiner Form bedeutsam für das persönliche Wohlbefinden sind, die Stressbewältigung dient der Aufrechterhaltung bzw. der Wiederherstellung dieses Wohlbefindens. Auf Seiten der Antezedentien kann zwischen Umwelt- bzw. Situationsmerkmalen (z.B. bestimmte Ressourcen wie die Größe des sozialen Netzwerks) und Personenmerkmalen (z.B. bestimmte Persönlichkeitseigenschaften oder Werte und Ziele einer Person) unterschieden werden. Schließlich wird auf Seiten der langfristigen Konsequenzen zwischen dem psychischen Wohlbefinden, der körperlichen Gesundheit und dem Sozialverhalten unterschieden. In Abbildung 1 ist angedeutet, dass der transaktionalen Stresstheorie eine strukturelle Annahme unterliegt. Dem Stressprozess gehen bestimmte Antezedentien voraus. Die Stressbewältigung wiederum hat sowohl kurz- als auch langfristige Konsequenzen. Gleichzeitig soll mit den rekursiven Pfeilen auf den prozesshaften Charakter

der transaktionalen Sichtweise von Stress hingewiesen werden. So wirkt die Stressbewältigung über ihre Veränderung der Person-Umwelt-Relation zurück auf die Bewertungsprozesse.

Die komplexen Beziehungen zwischen den Modellkomponenten werden weiter unten näher expliziert (Kap. 3.1.4). Zuvor sollen aber die einzelnen Komponenten der Stresstheorie erläutert werden.

3.1 *Bewertungsprozesse*

In dem Modell werden drei Arten der Bewertung unterschieden, nämlich die primäre Bewertung (Ereigniseinschätzung; „primary appraisal“), die sekundäre Bewertung (Ressourceneinschätzung; „secondary appraisal“) und die Neubewertung („reappraisal“), die jeweils unterschiedliche Funktionen haben und auf verschiedenartigen Informationsquellen basieren (Jerusalem, 1990; Lazarus & Folkman, 1984).

In der *Primärbewertung* wird die subjektive Bedeutsamkeit der situativen Anforderungen überprüft. Die Situation kann als irrelevant, angenehm-positiv oder als stressrelevant beurteilt werden (Lazarus & Launier, 1978; in der Veröffentlichung von 1966 unterscheidet Lazarus zwischen bedrohlich vs. nicht bedrohlich). Als irrelevant wird eine Situation dann eingeschätzt, wenn die Situation unbedeutsam ist und somit die Person sich nicht weiter mit der Situation auseinandersetzen muss. Angenehm-positive Einschätzungen treten auf, wenn persönliche Kompetenzen den gestellten Anforderungen als deutlich überlegen angesehen werden. Zu einer stressrelevanten Einschätzung kommt es, wenn die Situation und deren Bewältigung als persönlich wichtig empfunden werden und das Verhältnis zwischen Anforderungen und Bewältigungsmöglichkeiten unklar ist bzw. zu Ungunsten der Person ausfällt.

Weiterhin unterscheidet Lazarus drei stressrelevante Bedingungen: Schädigung/Verlust, Bedrohung und Herausforderung (ausführlich dazu Jerusalem, 1990; s.a. Lazarus & Folkman, 1984). Schädigung oder Verlust („harm/loss“) beinhaltet eine schon vorliegende Beeinträchtigung des Wohlergehens, die nicht mehr abwendbar ist. Bedrohung („threat“) bezieht sich auf die Antizipation einer Beeinträchtigung, also auf einen zukünftigen Schaden oder Verlust. Die Wahrnehmung einer Herausforderung („challenge“) ist wie die Bedrohung auf die Zukunft gerichtet, es steht jedoch ein möglicher Erfolg bzw. Gewinn in einer schwierigen Situation kognitiv im Vordergrund. Darüber hinaus unterscheiden sich Herausforderung und Bedrohung auch affektiv, da Herausforderung mit positiven und Bedrohung mit negativen Gefühlen einhergehen (vgl. die Unterscheidung von Distress und Eustress von Selye, 1976a). Die unterschiedlichen Formen primärer Bewertung sind nicht unabhängig voneinander, sondern kön-

nen in Abhängigkeit vom Ausmaß der Ambiguität einer Situation vermischt auftreten (Folkman & Lazarus, 1985; Stone & Neale, 1984).

Wird eine stressrelevante Situation als herausfordernd, bedrohlich und/oder schädigend beurteilt, wird die *Sekundärbewertung* salient. Der Begriff sekundär bedeutet nicht, dass dieser Bewertungsschritt nach der primären Bewertung erfolgen muss oder weniger wichtig ist. Vielmehr betont Lazarus (Lazarus & Folkman, 1984, 1987), dass die Bewertungsprozesse sich zeitlich überschneiden oder gemeinsam auftreten können. Auch denkbar sind Situationen, in denen die Sekundärbewertung zeitlich vorausgeht, wenn beispielsweise für bestimmte Situationen Bewältigungshandlungen kognitiv repräsentiert sind, die die Einschätzung von Bedrohung erst gar nicht aufkommen lassen. Jerusalem (1990, S. 12) gibt hierzu ein passendes Beispiel: „Man stelle sich einen Ehemann vor, der im Laufe vieler Jahre gelernt hat, daß bei später Heimkehr das Mitbringen eines Blumenstraußes die Wut seiner Ehefrau schnell verrauchen lässt.“

Das Individuum vollzieht in der Sekundärbewertung eine Abschätzung seiner Ressourcen und Möglichkeiten in Hinblick auf einen erfolgreichen Abschluss der stressbezogenen Auseinandersetzung. Lazarus und Folkman (1984, S. 35) definieren sekundäre Bewertungen als einen „...complex evaluative process that takes into account which coping options are available, the likelihood that a given coping option will accomplish what it is supposed to, and the likelihood that one can apply a particular strategy or set of strategies effectively.“ Diese Definition orientiert sich an Banduras' (1977, 1982) Unterscheidung zwischen der Ergebniserwartung und der Wirksamkeitserwartung. Die Ergebniserwartung bezieht sich auf die Einschätzung einer Person, ob ein gegebenes Verhalten zu einem bestimmten Ergebnis führt, die Wirksamkeitserwartung meint die Überzeugung einer Person, dass sie das Verhalten, welches benötigt wird, um das Ergebnis zu erreichen, erfolgreich ausführen kann.

Zu einer *Neubewertung* der Person-Umwelt-Beziehung kann es im Verlauf der Auseinandersetzung mit der Umwelt und den dadurch eventuell modifizierten situativen Bedingungen kommen (Lazarus & Folkman, 1984). Neubewertungen unterscheiden sich inhaltlich nicht von der primären und sekundären Bewertung, sie erfolgen lediglich zeitlich später, wenn zum Beispiel neue Informationen über die Person-Umwelt-Beziehung vorliegen (Jerusalem, 1990). Erneut schätzt eine Person ab, ob die veränderte Person-Umwelt-Beziehung das Wohlbefinden bedroht und ob ausreichende Ressourcen zur Bewältigung zur Verfügung stehen.

Die kognitiven Bewertungen sind der Theorie nach auf bestimmte Person-Umwelt-Relationen beschränkt und können zwischen Situationen variieren. Lazarus (1991a; Lazarus & Launier, 1978) schließt aber nicht aus, dass es auch bestimmte Bewertungsstile („appraisal

styles“) gibt, die sich in verschiedenen Stresssituationen wiederholen und somit interindividuelle Unterschiede widerspiegeln.

Beide Bewertungsprozesse sind multiplikativ miteinander verknüpft. Lazarus und Folkman (1984, S. 35, Hervorhebung durch den Autor; s.a. Jerusalem, 1990) beschreiben die Beziehung der beiden Bewertungsprozesse wie folgt: „Secondary appraisals of coping options and primary appraisals of what is at stake *interact* with each other in shaping the degree of stress and the strength and quality (or content) of the emotional reaction.“ So kann eine primäre Bewertung einer Situation als Bedrohung in ihrem Belastungsausmaß abgeschwächt werden durch eine positive Bewertung der Bewältigungsmöglichkeiten.

Primäre und sekundäre Bewertung gemeinsam bewirken eine Stressreaktion und haben schließlich einen Einfluss auf die (Bewältigungs-)Strategien, die eine Person einsetzt, um mit einer stressigen Situation umzugehen. Im Folgenden werden statt Primär- und Sekundärbewertung die Bezeichnungen Ereignis- und Ressourceneinschätzung verwendet, da sie der inhaltlichen Bedeutung der beiden Bewertungsformen besser entspricht (Schwarzer, 1998b, 2000).

3.2 *Stressbewältigung*

Der Begriff der Stressbewältigung umfasst nach Lazarus und Launier (1978) alle Anstrengungen einer Person, mit stressrelevanten Situationen umzugehen, mit dem Ziel, die Anforderungen zu mindern, zu meistern oder zu tolerieren (vgl. Folkman, Lazarus, Gruen & DeLongis, 1986). Allerdings ist diese Sichtweise von Bewältigung nicht unumstritten und eine einheitliche Definition ist nur schwer zu finden (Schwarz, Salewski & Tesch-Römer, 1997). Dennoch orientieren sich viele Bewältigungsforscher an der breiten Definition von Lazarus (vgl. Folkman & Moskowitz, 2004): So fasst Weber (1997, S. 7; s.a. auch Greve, 1997; Wentura, Greve & Klauer, 2002) zusammen, dass der Begriff der Bewältigung in verschiedenen Kontexten in dem Sinne verwendet wird, „...dass mit einer Situation umzugehen ist, die aus objektiver Sicht, d.h. auf der Grundlage eines intersubjektiven Konsenses bezüglich der Belastungshaftigkeit, und/oder aus subjektiver Sicht des Betroffenen in irgendeiner Weise belastend, schwierig, fordernd, unangenehm ist.“ Zusammenfassend lässt sich sagen, dass Bewältigung ein prozesshaftes Geschehen ist, das sich auf jede Form kognitiver und behavioraler Bemühungen einer Person bezieht, mit Ist-Soll Diskrepanzen umzugehen, die sich ihr hinsichtlich ihres Wohlbefindens und/oder ihrer Handlungsfähigkeit in einer spezifischen Person-Umwelt-Situation stellen (Laux & Weber, 1990; Lazarus, 1991a; Schwarzer & Schwarzer, 1996; Weber, 1997; Wentura, Greve & Klauer, 2002).

Auch die Klassifikation von Bewältigungsstrategien bzw. –prozessen ist problematisch und in der Literatur kontrovers diskutiert (vgl. Greve, 1997; Weber, 1997). Lazarus und seine Arbeitsgruppe (Cohen & Lazarus, 1979; Folkman, 1984; Lazarus, 1991a; Lazarus & Folkman, 1984; Lazarus & Launier, 1978) unterscheiden zwischen problem- und emotionszentrierten Reaktionen (*problem-focused vs. emotion-focused coping* oder instrumentelle vs. palliative Bewältigung). Die problemzentrierte Bewältigung hat die Funktion, die bewältigungsrelevante Konflikt- oder Problemstruktur zu verändern, während die emotionszentrierte Bewältigung negative Emotionen regulieren soll. Auch andere AutorInnen treffen eine ähnliche, auf zwei Kategorien basierende Unterscheidung (z.B. Krohne, 1989, 1994: kognitiv vermeidende vs. vigilante Reaktion; Brandstädter, 1989; Rothermund & Brandstädter, 1997: assimilative vs. akkomodative Bewältigung; Tobin, Holroyd, Reynolds & Wigal, 1989: engagement vs. disengagement; für einen aktuellen Überblick s. Folkman & Moskowitz, 2004; s. für andere mehrdimensionale Ansätze z.B. Klauer, Philipp & Ferring, 1989).

Lazarus (Folkman & Lazarus, 1985; Lazarus & Folkman, 1984) geht nicht davon aus, dass bestimmte Bewältigungsstrategien immer mit positiveren Ergebnissen einhergehen als andere. Je nach Situation können eher problemzentrierte oder aber emotionszentrierte Bewältigungsstrategien in einer besseren Anpassung an die stressige Situation resultieren. Eine Reihe von Untersuchungen zeigen dagegen, dass abwehrende und emotionszentrierte Bewältigungsstrategien mit negativeren Ergebnissen für Personen zusammenhängen (z.B. Aldwin & Revenson, 1987; Billings & Moos, 1984; Bolger, 1990; Carver et al., 1993; Felton, Revenson & Hinrichsen, 1984; Holohan & Moos, 1985; McCrae & Costa, 1986). Aber auch problemzentrierte Bewältigungsstrategien haben zum Teil negative Konsequenzen zur Folge (z.B. Bolger, 1990; Carver & Scheier, 1994; Mattlin, Wethington & Kessler, 1990; für positive Effekte von problemzentrierten Strategien s. z.B. Aldwin & Revenson, 1987; Aspinwall & Taylor, 1992; Glyshaw, Cohen & Towbes, 1989). Eine wichtige Rolle im Hinblick auf die Effektivität problem- vs. emotionszentrierter Bewältigungsstrategien wird der Kontrollierbarkeit von stressigen Situationen zugeschrieben. Folkman, Schaefer und Lazarus (1979) haben das Konzept des „goodness of fit“ hierfür eingeführt. Nach dem goodness of fit-Modell der Bewältigungseffektivität ist die Effektivität einer bestimmten Bewältigungsstrategie von der Übereinstimmung oder dem goodness of fit zwischen der Bewältigungsstrategie und der Kontrollierbarkeit einer Situation abhängig. Problemzentrierte Strategien sollen nach diesem Modell bei kontrollierbaren Ereignissen adaptiv sein, während emotionszentrierte Strategien in solchen Situationen maladaptiv sein sollen. Dagegen sind emotionszentrierte Bewältigungsstrategien dann effektiv, wenn ein Ereignis als unkontrollierbar bewertet wird. In sol-

chen Situationen sollen problemzentrierte Strategien maladaptiv sein. Die empirische Befundlage zu diesem Modell ist gemixt. Felton und Revenson (1984) wie auch Masel, Terry und Gribble (1996) finden keine Unterstützung. In beiden Untersuchungen hat die Kontrollierbarkeit eines Ereignisses keinen Effekt auf die Effektivität problem- vs. emotionszentrierter Bewältigungsstrategien. Problemzentrierte Strategien erwiesen sich unabhängig von der Kontrollierbarkeit einer Situation als effektiv, emotionszentrierte Strategien als ineffektiv. Forsythe und Compas (1987) und Vitaliano, DeWolfe, Maiuro, Russo und Kanton (1990) finden dagegen unterstützende Ergebnisse für die moderierende Rolle situationaler Kontrollbewertungen. In der Studie von Forsythe und Compas erwiesen sich problemzentrierte Strategien als effektiver im Umgang mit kontrollierbaren kritischen Lebensereignissen, während bei unkontrollierbaren Lebensereignissen sich emotionszentrierte Strategien als effektiver zeigten. Allerdings konnte dieses Ergebnismuster nicht bei alltäglichen, belastenden Situationen gefunden werden. Vitaliano et al. (1990) finden für problemzentrierte Strategien einen Modatoreffekt der Kontrollierbarkeit einer Situation. Für emotionszentrierte Strategien ließ sich dies allerdings nicht zeigen.

Auch die Messung von Bewältigungsstrategien wird in der Stress- und Bewältigungsforschung kontrovers diskutiert (vgl. Folkman & Moskowitz, 2004; Lazarus & Launier, 1978; Perrez, 1992). Zu den häufig eingesetzten Methoden gehören standardisierte Fragebogen, in denen die untersuchten Personen angeben sollen, in welchem Ausmaß sie bestimmte Gedanken oder Verhaltensweisen nutzen, um mit stressigen Situationen umzugehen. Bekannte Beispiele für solche Bewältigungsinventare sind international die Ways of Coping Checklist (Folkman & Lazarus, 1980), die COPE (Carver, Scheier & Weintraub, 1989), das Coping Responses Inventory (Moos, 1988), der Coping Strategy Indicator (Amirkhan, 1990) oder das Coping Inventory for Stressful Situations (Endler & Parker, 1990), deutsche Fragebogen sind z.B. das Angstbewältigungsinventar (Krohne, Schumacher & Egloff, 1992) oder die Trierer Skalen zur Krankheitsbewältigung (Klauer et al., 1989).

Schwarzer und Schwarzer (1996) geben einen ausführlichen Überblick über verschiedene Bewältigungsfragebogen. Generell muss bei der Erhebung von Bewältigungsstrategien die Stabilität von Bewältigung, also die Konsistenz von Bewältigungsantworten über Situationen hinweg und die Dimensionalität, auf die bereits oben eingegangen wurde, mit berücksichtigt werden. Die transaktionale Stresstheorie sieht Bewältigung als Prozess an. Es gibt aber auch theoretische Ansätze, die von stabilen Bewältigungsstilen ausgehen (z.B. „repressor vs. sensitizer“, Byrne, 1964; „monitoring vs. blunting“, Miller, 1980; Vigilanz vs. Vermeidung, Averill & Rosenn, 1972). Gängige Untersuchungen zur Bewältigung fokussieren meist auf inter-

individuellen Unterschieden, d.h. Unterschiede im Muster von Bewältigungsreaktionen zwischen Personen. Die Konsistenz von Bewältigung über Situationen hinweg bzw. die Generalität von Bewältigung bezieht sich wiederum darauf, ob Bewältigung von der Situation, der Person und der Interaktion (Transaktion) von Situation und Person abhängig ist. Wenn interindividuelle Unterschiede vorliegen, muss eine gewisse Stabilität in der Form der Bewältigung vorliegen (vgl. Folkman & Lazarus, 1981).

Schwarzer und Schwarzer (1996) empfehlen für die zukünftige Bewältigungsforschung eine Erfassung von Bewältigung auf mehreren Ebenen wie z.B. die gleichzeitige Erfassung von situationalen und stabilen Bewältigungsreaktionen bzw. Bewältigungsstilen (vgl. Carver & Scheier, 1994; Carver et al., 1989).

Im Folgenden werden die Begriffe Bewältigungsverhalten, Bewältigungsstrategien und Coping synonym verwendet, gemeint sind immer Anstrengungen eines Individuums, ob verhaltensmäßig oder kognitiv, mit einer stressigen Situation umzugehen.

3.3 Antezedentien und Ergebnisse des Stressprozesses

Die Bewertungsprozesse werden durch bestimmte Antezedentien beeinflusst. Die Bewertung und Bewältigung einer stressigen Situation haben bestimmte Ergebnisse zur Folge.

Auf Seiten der Antezedentien kann zwischen Umweltvariablen und Personenvariablen bzw. zwischen sozialen und personalen Ressourcen unterschieden werden (Lazarus, 1991a; Lazarus & Folkman, 1984; Moos & Billings, 1982; Schwarzer, 2000).

Lazarus und Folkman (1984, S. 55ff) nennen zwei wichtige personale Ressourcen: Zum einen sind dies „commitments“, die Motivationsstrukturen repräsentieren wie persönliche Ziele und Intentionen und somit die subjektive Bedeutsamkeit bestimmter Situationen beeinflussen. Zum anderen sind es „beliefs“, die Überzeugungen und Erwartungen einer Person reflektieren, bestimmten situativen Anforderungen gewachsen zu sein. Unter „beliefs“ fallen Konstrukte wie generalisierte Selbstwirksamkeitserwartungen, Kontrollüberzeugungen, Ängstlichkeit oder Selbstbewusstsein (Schwarzer, 1998b). Während Lazarus zunächst (Cohen & Lazarus, 1979; Folkman & Lazarus, 1980) stabilen Bewertungs- und Bewältigungsstilen, die auf bestimmte Persönlichkeitseigenschaften im Sinne von „Traits“ rückführbar sind, eine Bedeutung abspricht und die prozessuale Sichtweise von Stress betont, heben andere AutorInnen die Wichtigkeit von globaleren, übergeordneten Persönlichkeitseigenschaften hervor (vgl. Bolger & Zuckerman, 1995). Insbesondere Neurotizismus wird als kritische Variable angesehen, die einen Einfluss auf die kognitive Bewertung und den Bewältigungsprozess hat (z.B. Ben-Porath & Tellegen, 1990; McCrae, 1990). In neueren Publikationen weicht auch

Lazarus von seiner ursprünglichen Position ab. Neben dem prozessualen Charakter von Stress geht er nun auch von zeitlich stabilen Bewertungs- und Bewältigungsstilen aus (Lazarus, 1990, 1993a, 1999).

Unter Umweltvariablen fasst Lazarus (1991a; Lazarus & Folkman, 1984) formale Eigenschaften von Situationen wie Neuigkeit der Situation, Ungewissheit über den Eintritt von Ereignissen, Ambiguität und temporale Aspekte (z.B. Imminenz, Dauer) der stressigen Bedingung. Situationen, die schwierig, unvorhersagbar, andauernd und mehrdeutig sind, sollen eine höhere Bedrohungswahrnehmung auslösen als leichte Situationen, die erwartbar und nur von kurzer Dauer sind.

Der Stressprozess hat nach Lazarus (Lazarus & Folkman, 1984) Folgen für das Sozialverhalten, das psychische Wohlbefinden und die somatische Gesundheit. Weiterhin kann zwischen kurzfristigen und langfristigen Folgen unterschieden werden. Diese Unterscheidung ist insofern wichtig, da kurzfristige und langfristige Folgen nicht übereinstimmen müssen (Moos, Ralph & Swindle, 1990). So ist es möglich, dass eine bestimmte Person-Umwelt-Transaktion kurzfristig erfolgreich bewältigt werden kann, langfristig aber durchaus auch negative Konsequenzen für die oben benannten Bereiche folgen (z.B. Aspinwall & Taylor, 1992).

Unter Sozialverhalten („social functioning“) fasst Lazarus (Lazarus & Folkman, 1984) die Zufriedenheit mit interpersonalen Beziehungen, das Ausmaß, in dem soziale Rollen erfüllt werden, und die notwendigen Fähigkeiten, die sozialen Rollen und interpersonalen Beziehungen aufrecht zu erhalten. All diese Formen des Sozialverhaltens hängen von der Effektivität ab, mit der man alltägliche Situationen bewertet und bewältigt. Sowohl kurzfristig kann das Sozialverhalten beeinträchtigt sein als auch langfristig, wenn die Adaption an chronische, problematische Person-Umwelt-Transaktionen nicht gelingt.

Hinsichtlich des psychischen Wohlbefindens kann kurzfristig zwischen spezifischen Emotionen in einer gegebenen Situation und langfristig zwischen dem generellen psychischen Wohlbefinden unterschieden werden.

Schließlich kann auf Seiten der somatischen Gesundheit zwischen physiologischen Reaktionsmustern, die in bestimmten Situationen ausgelöst werden, und eher längerfristigen Konsequenzen, die sich in bestimmten körperlichen Symptomen oder auch Krankheiten niederschlagen, differenziert werden.

Das Sozialverhalten, das psychische Wohlbefinden und die somatische Gesundheit stehen untereinander in Beziehung und sind möglichst bei der Bewertung des Ergebnisses eines Stressprozesses zu berücksichtigen. So kann es sein, dass in einem Bereich ein positives (z.B. wenig körperliche Beschwerden), in einem anderen Bereich aber ein negatives Ergebnis

(z.B. gestörte soziale Beziehungen) als Folge des Umgangs mit einer bestimmten Situation auftritt. Somit sind auch für die Bereiche ähnlich wie für kurz- und langfristige Effekte gegenläufige Ergebnisse möglich.

3.4 Wirkungszusammenhänge der Modellkomponenten

Die transaktionale Stresstheorie macht strukturelle Annahmen über die Beziehung der in Abbildung 1 skizzierten Modellkomponenten, gleichzeitig ist nach dieser transaktionalen Konzeption Stress ein prozesshaftes Geschehen. Die Person-Umwelt-Relation verändert sich über die Zeit, was auf die Bewertung und die Bewältigung einer Situation rückwirkt. Folgend sollen diese komplexen Wirkungszusammenhänge, wie sie in Abbildung 1 angedeutet sind, zusammenfassend dargestellt werden.

Zunächst scheint der transaktionalen Stresstheorie allein eine lineare, deterministische Wirkkette zu Grunde zu liegen. Bestimmte Umwelt- und Personenvariablen beeinflussen die kognitive Bewertung (Ereignis- und Ressourceneinschätzung) eines Stressors. In Abhängigkeit der kognitiven Bewertung werden bestimmte Bewältigungsstrategien ausgewählt, mit dem Ziel, eine gestörte Person-Umwelt-Relation zu ändern. Die Bewältigung hat schließlich sowohl kurz- als auch langfristige Effekte auf das psychische Wohlbefinden, die körperliche Gesundheit und das Sozialverhalten einer Person. Der Stressprozess ist nach der transaktionalen Stresstheorie aber durch eine starke Dynamik gekennzeichnet (Lazarus & Launier, 1978). Die Person wirkt auch auf die Umwelt ein. Bewältigung verändert die Person-Umwelt-Beziehung, wodurch es zu einer Neubewertung der Situation bzw. des Stressors kommen kann, was wiederum einen Einfluss auf das Bewältigungsverhalten haben kann. Somit unterliegt den Modellkomponenten nicht eine rein lineare Wirkkette, sondern die Dynamik des Stressprozesses zeichnet sich durch Rückkopplungen zwischen den Komponenten aus. Schwarzer (1998a) kommt basierend auf dem komplexen und transaktionalen Charakter zu dem Schluss, dass die transaktionale Stresstheorie als ganzes Modell nicht empirisch überprüfbar sei. Wissenschaftliche Untersuchungen müssten bei der Auswahl adäquater Untersuchungsmethoden nach Kompromissen dahingehend suchen, welche Ausschnitte sie aus dem Gesamtmodell untersuchen möchten. Allerdings betont er, dass Untersuchungen nicht alleine auf die Struktur fokussieren, sondern auch den Prozess berücksichtigen sollen. Er fordert, dass „...stress appraisal processes need to be predicted by environmental and personal variables as antecedents, and coping strategies and long-term affects need to be considered“ (S. 532). Im folgenden Kapitel werden Kriterien präzisiert, die empirische Untersuchungen erfüllen müssen, um eine adäquate Überprüfung der transaktionalen Annahmen zu ermögli-

chen. Sowohl die strukturellen Annahmen wie auch der Stressprozess sollten bei einer Analyse berücksichtigt werden. Bisherige empirische Untersuchungen mit dem Ziel der Überprüfung der transaktionalen Stresstheorie werden in Hinblick auf diese Kriterien begutachtet. Zunächst soll aber eine prominente Weiterentwicklung der transaktionalen Stresstheorie vorgestellt werden. Diese Theorie steht nicht im Fokus dieser Arbeit, soll aber der Vollständigkeit wegen kurz beschrieben werden, um auch aktuellen Tendenzen gerecht zu werden.

3.5 *Weiterentwicklung der transaktionalen Stresstheorie*

Eine Weiterentwicklung der transaktionalen Stresstheorie stellt die Theorie der Ressourcenerhaltung von Hobfoll (1988, 1989, 1998, 2001; vgl. Schwarzer, 2001) dar. Auch Hobfoll geht von einer kognitiv-transaktionalen Sichtweise von Stress aus, allerdings stehen in seiner Theorie die Stressbewältigung und deren Gründe im Mittelpunkt, während bei Lazarus die kognitive Bewertung im Vordergrund steht. Zentral für den Ansatz von Hobfoll ist die Annahme, dass Menschen nach der Erhaltung, dem Erwerb und der Stärkung von Ressourcen streben. Stress entsteht dann, wenn der Verlust von Ressourcen droht, Ressourcen tatsächlich verloren sind oder keine neuen Ressourcen gewonnen werden können, nachdem bestimmte Ressourcen investiert wurden. Auch in der Theorie von Lazarus spielen Ressourcen eine wichtige Rolle (personale und soziale Ressourcen, s.o.), allerdings liegt der Unterschied zu Hobfoll darin, dass Lazarus auf subjektive Ressourcen (wie z.B. generalisierte Selbstwirksamkeitserwartungen) fokussiert, während Hobfoll objektiven Ressourcen ein weitaus größeres Gewicht beimisst. Objektive Ressourcen sind nach Hobfoll Gegenstände (z.B. Nahrungsmittel), Bedingungen (z.B. berufliche Karriere), Persönlichkeitsmerkmale (z.B. persönliche Überzeugungen) und Energien (z.B. Geld). Der Verlust solcher Ressourcen soll nach Hobfoll Stress auslösen.

Die Theorie der Ressourcenerhaltung stellt sicherlich eine wichtige Weiterentwicklung der transaktionalen Stresstheorie von Lazarus dar, da die Wichtigkeit von Kontextfaktoren, kulturellen Skripten und Mechanismen des Verlusts und Erwerbs von Ressourcen betont wird, was in der transaktionalen Sichtweise von Lazarus nur wenig Berücksichtigung findet. Dennoch sind Bewertungsprozesse und auch die Bewältigung von stressigen Situationen wichtige proximale Determinanten von Stress. Beide Ansätze sollten sich in Zukunft ergänzen und gegenseitig befruchten (vgl. Thompson & Cooper, 2001). Kontextfaktoren spielen in der vorliegenden Arbeit keine Rolle, auf die Bedeutsamkeit dieser Faktoren soll aber hingewiesen werden. So kann beispielsweise Weber (2003) den Einfluss von sozialen Normen auf das

Bewältigungsverhalten zeigen. Normabweichendes Bewältigungsverhalten wird missbilligt, was sich wiederum auf die Auswahl von Bewältigungsstrategien auswirken kann.

4 Die empirische Überprüfung der transaktionalen Stresstheorie

In diesem Kapitel sollen ausgewählte Befunde zur transaktionalen Stresstheorie vorgestellt werden. Basierend auf oder zumindest orientiert an der Theorie von Lazarus sind eine Vielzahl von Untersuchungen in unterschiedlichsten Bereichen durchgeführt worden, z.B. zur Erklärung der Bewältigung chronischer Erkrankungen (Maes, Leventhal & de Ridder, 1996) und bei beruflichem Stress (Crandall & Perrewé, 1995). Es ist unmöglich, einen umfassenden Überblick über die Empirie zur transaktionalen Stresstheorie zu geben, letztlich können nur Ausschnitte aufgelistet und bewertet werden. Ziel der vorliegenden Arbeit ist die empirische Überprüfung der postulierten Struktur der Kernvariablen im Rahmen der transaktionalen Stresstheorie, so wie sie in Abbildung 1 skizziert ist und den prozessualen Annahmen. Zunächst werden Kriterien beschrieben, die empirische Untersuchungen, die als Ziel die Überprüfung der transaktionalen Stresstheorie haben, erfüllen sollten (Kapitel 4.1). Hieran anschließend werden bisherige Untersuchungen gesichtet (Kap. 4.2) und schließlich im Hinblick auf die aufgestellten Kriterien kritisch bewertet (Kap. 4.3).

4.1 Kriterien für eine empirische Überprüfung der transaktionalen Stresstheorie

Die transaktionale Stresstheorie mit ihrer komplexen Struktur stellt an empirische Untersuchungen eine Reihe von Anforderungen: Zunächst wird aus Abbildung 1 deutlich, dass der Theorie ein komplexes Wirkungsgefüge unterliegt. Zentral sind die kognitiven Bewertungen, bestehend aus der Ereignis- und der Ressourceneinschätzung, die wiederum gemeinsam das Bewältigungsverhalten bedingen. Die Art und Weise, wie eine Person mit der stressigen Situation umgeht, hat zum einen kurz- und langfristige Konsequenzen, gleichzeitig wirkt dies auch auf die kognitiven Bewertungen und damit auch auf das Bewältigungsverhalten zurück. Wichtig für diesen Stressprozess sind darüber hinaus situative Faktoren, wie z.B. die Kontrollierbarkeit einer Situation, und personale Ressourcen, die beide einen Einfluss auf die kognitive Bewertung haben. Dieses komplexe Wirkungsgeflecht hat direkte Konsequenzen für die Überprüfung der Theorie. Dies sind zum einen die Wirkungsrichtung der Modellvariablen, also die Struktur, die den Theoriekomponenten unterliegt und zum anderen der der Theorie inhärente Prozess. Diese grundlegenden Annahmen erfordern für empirische Studien und deren statistische Auswertungsverfahren zweierlei Dinge: (a) Das Wirkungsgeflecht kann prinzipiell nur in experimentellen oder Längsschnittuntersuchungen überprüft werden; (b) Zur Prüfung der Modellannahmen müssen sowohl interindividuelle Unterschiede (z.B. Unterschiede im Ausmaß der personalen Ressourcen) als auch intraindividuelle Veränderungen

(z.B. Veränderung des psychischen Wohlbefindens im Laufe der Auseinandersetzung mit einem Stressor) analysiert werden. Dabei basiert die Struktur der Stresstheorie primär auf interindividuellen Unterschieden, während der Stressprozess durch intraindividuelle Veränderungen über die Zeit abgebildet werden kann.

Sowohl die Struktur als auch der Prozess müssen idealerweise gleichzeitig analysiert werden. Die hier aufgeführten Kriterien dienen als Grundlage für die Bewertung von empirischen Untersuchungen, die eine Überprüfung der transaktionalen Stresstheorie vorgenommen haben.

4.2 Die transaktionale Stresstheorie in ihrer empirischen Bewährung

Wie in der Einleitung konstatiert, gibt es eine kaum zu überblickende Anzahl an Studien, die sich mit Stress und dessen Bewältigung im weitesten Sinne beschäftigt haben (Hobfoll, Schwarzer & Chon, 1998; Somerfield & McCrae, 2000). Ein nicht unerheblicher Teil dieser Studien hat als theoretische Grundlage die transaktionale Stresstheorie. Eine genauere Inspektion der Studien zeigt aber, dass nur selten die transaktionale Stresstheorie in ihrer Gesamtheit Gegenstand der Prüfung war. Im vorliegenden Kapitel sollen Studien gesichtet werden, die die Vorhersagen und Struktur der Theorie geprüft haben. Es wird nicht der Anspruch einer vollständigen Auflistung erhoben. Vielmehr soll eine repräsentative Stichprobe dargestellt werden, um basierend auf einer kritischen Bewertung dieser Studien, der die oben beschriebenen Kriterien zu Grunde gelegt sind, zum Ziel und den Fragestellungen der vorliegenden Arbeit in Kapitel 5 zu gelangen.

Die gesichteten Studien lassen sich einerseits danach unterteilen, ob es (a) direktes Ziel war, die transaktionale Stresstheorie zu prüfen, und andererseits (b) nach methodischen Gesichtspunkten. Punkt 1 ergibt sich aus der Tatsache, dass auch Studien herangezogen werden können, deren Ziel nicht dezidiert die Überprüfung der Theorie war, die aber als indirekte Prüfung gelten können. Methodische Gesichtspunkte spielen eine noch bedeutendere Rolle, da den strukturellen und prozessualen Annahmen der Stresstheorie nicht alle gängigen Auswertungsverfahren und Untersuchungsdesigns gerecht werden. Die Bewertung methodischer Gesichtspunkte ist aus diesem Grund mitentscheidend für die Bewertung von empirischen Arbeiten. Auf Basis methodischer Gesichtspunkte lassen sich die ausgewählten Untersuchungen in sechs Kategorien unterteilen, deren Anspruch nicht absolute Trennschärfe ist, sondern gängige Untersuchungs- und Analyseansätze veranschaulichen sollen:

- 1) Die erste Kategorie bilden Studien, deren Fokus auf der intraindividuellen Ebene liegen, um Veränderungen in zentralen Modellvariablen über die Zeit abzubilden. Hierunter fal-

len Untersuchungen, die explizit den Stressprozess, also intraindividuelle Veränderungen in den zentralen Modellkomponenten, untersucht haben. Strukturelle Zusammenhänge werden nicht analysiert.

- 2) In die zweite Kategorie fallen korrelative Studien, die den Zusammenhang zwischen den verschiedenen Theoriebausteinen analysieren. Weder die komplexe Struktur, so wie sie in Abbildung 1 dargestellt ist, noch der Prozess werden weiter analysiert.
- 3) Kategorie 3 umfasst die so genannten „Examens-Studien“, die klassisch in der Stressforschung sind und aus diesem Grund eine eigene Kategorie bilden, auch wenn eine eindeutige Abgrenzung zu anderen Kategorien nicht möglich ist. Den Studien, die in diese Kategorie eingeordnet wurden, ist gemeinsam, dass sie den transaktionalen Stressprozess im Rahmen von bestimmten Situationen, meist Examina im Studium, untersucht haben. Die Analysen konzentrieren sich entweder auf intraindividuelle Veränderungen in den Modellkomponenten über die Zeit (und würden demnach auch in Kategorie 1 einordbar sein) oder sie analysieren den Zusammenhang der Modellkomponenten über die Zeit (und könnten auch der zweiten Kategorie zugeordnet werden).
- 4) Die vierte Kategorie beinhaltet experimentelle Untersuchungen, die bestimmte Konstrukte im Labor manipuliert haben, um so die Vorhersagen der Theorie zu prüfen. Experimentelle Untersuchungen ermöglichen es, kausale Annahmen zu prüfen und rekursive Beziehungen zu analysieren. Die Komplexität der transaktionalen Stresstheorie zwingt allerdings den experimentellen Ansatz, sich auf Ausschnitte zu konzentrieren.
- 5) Kategorie 5 subsumiert Studien, die zur Auswertung auf Pfad- und Strukturgleichungsanalysen zurückgreifen, um so die Struktur der Stresstheorie zu überprüfen. Pfad- und Strukturgleichungsanalysen (MacCallum & Austin, 2000; Marcoulides & Schumacker, 2001; Ulman & Bentler, 2003) ermöglichen es im Gegensatz zu einfachen Korrelations- oder Regressionsanalysen (s. Kategorie 2) komplexere Wirkstrukturen zu prüfen. Eine theoretisch angenommene Struktur wird auf ihre Passung auf Daten hin analysiert. Anhand statistischer Prüfwerte kann entschieden werden, ob die angenommenen Strukturzusammenhänge sich empirisch bestätigen lassen (in Kap. 6.2.1 wird das Verfahren ausführlich beschrieben).
- 6) In Kategorie 6 schließlich fällt eine Untersuchung, die gemäß oben aufgeführten Kriterien inter- und intraindividuelle Analyseverfahren kombiniert, die sowohl intraindividuelle Veränderungen über die Zeit als auch interindividuelle Unterschiede abbilden können. Keine der anderen Studien, die in die übrigen Kategorien fallen, haben beide Varianzquellen gleichzeitig analysiert.

Basierend auf dieser Kategorisierung werden die Studien im Folgenden dargestellt. Es wird jeweils auch angegeben, ob direktes Ziel die Überprüfung der Stresstheorie war oder die Untersuchung nur als indirekte Überprüfung gelten kann.

4.2.1 Abbildung intraindividuelle Veränderungen

In diese Kategorie fällt zum Beispiel eine Untersuchung aus der so genannten Berkeley-Gruppe um Richard S. Lazarus, die in den 80er Jahren des letzten Jahrhunderts ein groß angelegtes Forschungsprojekt zu Stress und dessen Bewältigung durchgeführt hat und deren primäres Ziel die Überprüfung der transaktionalen Stresstheorie war. Folkman, Lazarus, Dunkel-Schetter, DeLongis und Gruen (1986) untersuchten an 85 verheirateten Paaren mit mindestens einem Kind das Zusammenspiel der kognitiven Bewertung und dem Bewältigungsverhalten (in die Analysen gingen nur die Daten von 75 Paaren ein). Die AutorInnen verfolgten einen intraindividuellen Ansatz: Sie interviewten die Partner jeweils getrennt voneinander über einen Zeitraum von 6 Monaten insgesamt sechs Mal. Datenbasis für die Untersuchung bildeten die letzten fünf Interviews. Mittels strukturierter Protokolle wurden Angaben zur Ereignis- und Ressourceneinschätzung, zum Bewältigungsverhalten und zum Ergebnis des Stressprozesses (im Sinne einer Lösung der Situation, was als kurzfristiges Ergebnis gelten kann) für die in der zurückliegenden Zeit stressigste Situation erfasst. Zur Auswertung wurden die gesammelten Situationen einer Person jeweils aufgeteilt in Situationen, die über, und Situationen, die unter dem intraindividuellen Mittelwert für die Ereignis- und Ressourceneinschätzung sowie der Zufriedenheit der Befragten mit dem Ergebnis der Situation lagen. In Varianzanalysen wurden diese Situationen anschließend verglichen hinsichtlich der verwendeten Bewältigungsstrategien. Die Analysen zeigen, dass die Variabilität in der Nutzung bestimmter Bewältigungsstrategien in Teilen von der Ereignis- und Ressourceneinschätzung abhängt und bestimmte Bewältigungsstrategien wiederum mit einer unterschiedlichen Ergebniszufriedenheit zusammenhängen. Ein Problem der Analysen liegt darin, dass der Einfluss der Variablen jeweils separat voneinander analysiert und die mediierende Rolle der Bewältigungsstrategien nicht überprüft wurde. Auch die Rolle von Ressourcen wurde nicht weiter untersucht. Schließlich, was die AutorInnen auch selbst kritisch anmerken, sind die Analysen nur querschnittlicher Natur. Die Analysen basieren auf einem Vergleich zwischen Situationen. Für jede Situation wurden jeweils Angaben zu den Modellkomponenten retrospektiv erhoben. Kausale wie auch komplexere rekursive Beziehungen konnten nicht untersucht werden. Ein weiteres Problem ist, dass die stressigen Situationen zwischen den Personen variier-

ten und somit bestimmte Charakteristiken der Situationen (z.B. Kontrollierbarkeit, Dauer) nicht kontrolliert werden konnten.

4.2.2 Korrelative Untersuchungen

Unter korrelativen Untersuchungen werden solche Studien subsumiert, deren Auswertung auf einfachen korrelations- oder regressionsanalytischen Verfahren basiert. Hierunter fallen auch Längsschnittuntersuchungen.

Folkman, Lazarus, Gruen und DeLongis (1986) nutzten den gleichen Datensatz, auf dem auch die oben vorgestellte Studie basiert (Folkman, Lazarus, Dunkel-Schetter et al., 1986), um nun nicht mehr intraindividuelle Unterschiede zu analysieren, sondern um die Frage zu klären, welchen Einfluss bestimmte Persönlichkeitsmerkmale („Mastery beliefs“ und interpersonales Vertrauen), die Ereignis- und Ressourceneinschätzung und das Bewältigungsverhalten auf langfristige Ergebnisse des Stressprozesses haben. Langfristige Folgen wurden operationalisiert mit Maßen für das Ausmaß an psychischen Symptomen und der somatischen Gesundheit. Eine weitere Fragestellung fokussierte auf die Stabilität der Bewertungsprozesse und des Bewältigungsverhaltens über verschiedene Situationen hinweg. Die Autokorrelationen für die Bewältigungsdimensionen sind höher als für die kognitiven Bewertungen, wobei insgesamt die Autokorrelationen gering bis moderat ausfallen, was auf die situative Abhängigkeit beider Konstrukte hinweist (zum Teil kann dieser Effekt auch auf die geringe Reliabilität der Messverfahren zurückgeführt werden). Hierarchische multiple Regressionsanalysen, in denen die beiden Persönlichkeitsmerkmale, die kognitiven Bewertungen und das Bewältigungsverhalten Prädiktoren waren, zeigen, dass nur in dem Kriterium psychische Symptome signifikante Varianzanteile aufgeklärt werden konnten, wobei alle drei Konstruktbereiche, also Persönlichkeitsfaktoren, kognitive Bewertungen und Bewältigungsverhalten jeweils für sich signifikante Varianzanteile binden. Dagegen konnte die Varianz in den somatischen Symptomen nicht mit den Modellvariablen erklärt werden. Erklärung der Autoren ist, dass nur stabile Muster von kognitiven Bewertungen und Bewältigungsverhalten einen Einfluss auf die somatische Gesundheit haben können und dies somit eine Erklärung dafür sein könnte, warum in der Studie keine Varianz in körperlichen Symptomen aufgeklärt werden konnte. Insgesamt sprechen die Befunde dafür, dass Persönlichkeitsmerkmale wie auch kognitive Bewertungen und Bewältigungsverhalten einen Einfluss auf längerfristige Stressergebnisse haben. Aber auch in dieser Studie wurde das Modell mit seinen mediierenden Annahmen nicht direkt getestet.

Auch eine australische ForscherInnengruppe um Deborah J. Terry überprüfte in einer Reihe von Studien die transaktionale Stresstheorie von Lazarus (Terry, 1991, 1994; Terry, Tonge & Callan, 1995). In der Untersuchung von 1991 untersuchte Terry, ob situationale Bewertungen (Bedrohungsbewertung, Selbstwirksamkeitserwartungen, situative Kontrollüberzeugungen, Wichtigkeitsattributionen) die Effekte von personalen Ressourcen (Internale Kontrollüberzeugungen, Selbstwertschätzung, Neurotizismus, Leugnung, Ausmaß an sozialer Unterstützung) auf das Bewältigungsverhalten mediiieren. Untersuchungskontext bildete eine Prüfung im Rahmen eines Studiums (vgl. Kap. 4.2.3), Angaben der Studierenden zu den persönlichen Ressourcen und den situationalen Bewertungen wurden vier Wochen vor dem Examen, Angaben zum Bewältigungsverhalten direkt nach dem Examen erhoben. Sowohl die personalen Ressourcen als auch die situationalen Bewertungen haben einen Einfluss auf das Bewältigungsverhalten. Zum Beispiel hat soziale Unterstützung einen positiven Einfluss auf die Bewältigungsstrategie „Suche nach sozialer Unterstützung“ und einen negativen Effekt auf „Minimierung“, Selbstwirksamkeitserwartungen haben einen positiven Effekt auf „Instrumentelle Handlungen“ und einen negativen auf „Flucht/Selbstbeschuldigung“. Allerdings zeigen sich nicht die erwarteten Mediationseffekte der situationalen Bewertung. Nur wenige Effekte der personalen Ressourcen werden durch die situationalen Bewertungen mediiert. Dagegen konnten einige signifikante Interaktionseffekte zwischen den personalen Ressourcen und den situativen Bewertungen gefunden werden, die auf eine Pufferwirkung der Ressourcen hinweisen und somit gegen die transaktionale Stresstheorie sprechen, die von einer mediiierenden Wirkung der situativen Bewertungen ausgeht. Die regressionsanalytische Auswertung erlaubt allerdings keine komplexe Analyse des kausalen Zusammenhangsmusters zwischen den einzelnen Bausteinen der Stresstheorie.

In einer weiteren Untersuchung mit drei Messzeitpunkten konnte Terry (1994) dagegen die mediiierenden Effekte der situationalen Bewertung bestätigen. Zum ersten Messzeitpunkt wurden Angaben zu personalen Ressourcen (Internale Kontrollüberzeugungen, Selbstwertschätzung, Neurotizismus, Leugnung, Typ A-Verhalten, Ausmaß an sozialer Unterstützung), zum zweiten Messzeitpunkt verschiedene Maße für die situative Bewertung (Ausmaß an Stress, situationsspezifische Kontrollüberzeugungen, Selbstwirksamkeitserwartungen) erhoben. Das Bewältigungsverhalten wurde zu allen Messzeitpunkten erfasst. UntersuchungsteilnehmerInnen waren Studierende, zum ersten Messzeitpunkt sollten sie zusätzlich angeben, wie sie mit der stressigsten Situation der vergangenen vier Wochen umgegangen sind, zum zweiten Messzeitpunkt sollten sie die stressigste Situation beschreiben, mit der sie im Moment konfrontiert sind. Wie sie mit dieser Situation umgegangen sind (Bewältigungsver-

halten), wurde zwei Wochen später zum dritten Messzeitpunkt erfasst. Ein weiteres Ziel der Untersuchung war die Klärung der Frage, ob das Bewältigungsverhalten über Situationen hinweg stabil ist, d.h. ob früheres Verhalten späteres Verhalten vorhersagen kann. Die Analysen zeigen, dass situative Bewertungen (allerdings nur situationsspezifische Selbstwirksamkeitserwartungen) die Effekte von personalen Ressourcen auf das Bewältigungsverhalten mediiieren. Dieses Ergebnis stützt das theoretische Modell von Lazarus und kontrastiert die Ergebnisse der zuvor beschriebenen Untersuchung. Weiterhin lässt sich auch die Stabilität im Bewältigungsverhalten über verschiedene Situationen hinweg feststellen. Früheres Bewältigungsverhalten erklärt einen inkrementellen Anteil an Varianz an späterem Bewältigungsverhalten auf (bei situationaler Konsistenz, also sich ähnelnden stressigen Situationen, fielen die Zusammenhänge noch etwas höher aus), auch wenn personale Ressourcen und situative Bewertungen kontrolliert wurden. Darüber hinaus können auch die personalen Ressourcen und die situativen Bewertungen einen signifikanten Varianzanteil im Bewältigungsverhalten aufklären. Auch in dieser Studie wurden keine Ergebnismaße wie psychisches Wohlbefinden oder somatische Beschwerden erhoben, das komplexe Zusammenspiel zwischen den verschiedenen Theoriebausteinen wurde ebenfalls nicht berücksichtigt.

Auch zwei Untersuchungen von Kenneth I. Pakenham (1999; Pakenham & Rinaldis, 2001) zum Umgang mit und zur Anpassung an chronische Erkrankungen können zu den typischen Vertretern von korrelativen Untersuchungen gezählt werden. Beide Studien haben nicht das primäre Ziel, die transaktionale Stresstheorie zu überprüfen, die aber die theoretische Grundlage beider Untersuchungen bildete. Ziel war es, auf Basis der Theorie Unterschiede im Umgang mit und der daraus resultierenden Anpassung an eine chronische Erkrankung zu erklären.

Pakenham (1999) untersuchte in einer Längsschnittstudie mit Multiple Sklerose-Patienten den Einfluss von personalen Ressourcen (sozialer Unterstützung), der kognitiven Bewertung (Ereignis- und Ressourceneinschätzung) und dem Bewältigungsverhalten, die alle zum ersten Messzeitpunkt erhoben wurden, auf verschiedene Anpassungsindizes (globales Wohlbefinden, Depression, subjektiver Gesundheitszustand, etc.), die zum ersten Messzeitpunkt und 12 Monate später erneut erfasst wurden. Die Auswertung erfolgte über hierarchische Regressionsanalysen. Die Querschnittsanalyse (nur Messzeitpunkt 1) zeigt, dass sowohl die kognitive Bewertung als auch das Bewältigungsverhalten Einfluss auf die Anpassungsindizes aufweisen, soziale Unterstützung hat dagegen keinen eigenständigen Effekt. Die Längsschnittanalyse, in der die Anpassungsindizes zum ersten Messzeitpunkt kontrolliert

wurden, zeigt lediglich einen signifikanten Effekt von den Bewältigungsstrategien auf die Anpassungsindizes.

In einer Untersuchung mit AIDS-Patienten konnten Pakenham und Rinaldis (2001) vergleichbare Ergebnisse ermitteln. Auch hier zeigen hierarchische Regressionsanalysen, dass personale Ressourcen wie soziale Unterstützung, die kognitive Bewertung und das Bewältigungsverhalten signifikante Varianzanteile in verschiedenen Anpassungsindizes, wie depressive Symptome und allgemeines Wohlbefinden, aufklären können. Eine Überprüfung der strukturellen Annahmen wurde in keiner der beiden Studien vorgenommen, somit können die Ergebnisse nur als Hinweise auf die Bewährung der Theorie gedeutet werden

Auch eine Studie von Chang (1998) kann als indirekte Prüfung der transaktionalen Stresstheorie gewertet werden. Chang ging der Frage nach, ob die Effekte von dispositionellem Optimismus als relativ stabiles Persönlichkeitsmerkmal (vgl. Scheier & Carver, 1985) auf Konsequenzen von Stress über die kognitive Bewertung des Stressors und das Bewältigungsverhalten mediiert werden. Die Untersuchungsstichprobe setzte sich wiederum aus Studierenden zusammen, die am Tag eines Examens, welches als Zielstressor diente, den Fragebogen ausfüllen sollten. Es wurden neben dispositionellem Optimismus Angaben zur kognitiven Bewertung, Bewältigungsverhalten, depressiven Symptomen und somatischen Beschwerden erhoben. Hierarchische Regressionsanalysen zeigen, dass die Ereignis- und Ressourceneinschätzung einen signifikanten Anteil an Varianz in dem Bewältigungsverhalten aufklären kann, welches wiederum bedeutsame Effekte auf die Anpassungsindizes zeigt. Wie vorhergesagt und in Übereinstimmung mit den Annahmen der transaktionalen Stresstheorie wird ein Teil der Effekte von dispositionellem Optimismus auf die Anpassungsindizes durch die kognitive Bewertung und das Bewältigungsverhalten mediiert. Allerdings bleiben auch direkte Effekte bestehen.

Ein großes Manko einfacher korrelativer Studien generell ist, dass komplexe Wirkungszusammenhänge, wie in der transaktionalen Stresstheorie angenommen, nicht überprüft werden. Zwar können mit Längsschnittuntersuchungen, wie denen von Terry (1991, 1994) und Pakenham (1999; Pakenham & Rinaldis, 2001) Hinweise auf die kausale Beziehung der Modellvariablen gewonnen werden (z.B. Bewältigungsverhalten zu Messzeitpunkt 1 hängt mit Wohlbefinden zu Messzeitpunkt 2 zusammen), aber die Prüfung der vermittelnden Prozesse kann mit einfachen regressionsanalytischen Verfahren nicht vorgenommen werden.

4.2.3 „Examens-Studien“

Untersuchungen zum Umgang von Personen mit Stress im Allgemeinen und somit auch Untersuchungen, die die Überprüfung der transaktionalen Stresstheorie zum Ziel haben, sind häufig mit zwei grundsätzlichen Problemen konfrontiert. In vielen retrospektiven Befragungen sollen die Versuchspersonen sich an eine besonders stressige Situation erinnern. Zu dieser Situation werden dann Angaben zu der Bewertung und der Bewältigung erhoben. Zum einen unterscheiden sich die Stressoren zwischen Personen, wodurch bestimmte Charakteristiken dieser Ereignisse nicht kontrolliert werden können. Zum anderen soll der Prozess der Stressbewältigung, also der zeitliche Aspekt der stressbehafteten Transaktion zwischen Person und Umwelt, analysiert werden. Um beiden Problemen vorzubeugen, wurden u.a. Examina, denen sich Studierende konfrontiert sehen, als stressige Situation verwendet und entsprechend Studierende vor und nach einem Examen zu den Theoriebausteinen befragt. Examina bieten sich, neben der leichten Erreichbarkeit von Studierenden, an, da für alle Befragten die gleiche Situation vorliegt und durch die Vor- und Nachherbefragung der Stressprozess abgebildet werden kann.

Folkman und Lazarus (1985) waren die ersten, die ein Examen als Stressor nutzten, um die transaktionale Theorie zu prüfen. Die Studierenden wurden zu drei Zeitpunkten untersucht: Zwei Tage vor dem Examen (Messzeitpunkt 1), fünf Tage nach dem Examen und zwei Tage vor Bekanntgabe der Noten (Messzeitpunkt 2) und fünf Tage nach Bekanntgabe der Noten (Messzeitpunkt 3). Zu jedem Messzeitpunkt wurden Angaben zur kognitiven Bewertung, zu Emotionen und zum Bewältigungsverhalten erhoben. Die Ergebnisse zeigen zunächst, dass theoriekonform die Emotionen, die in diesem Fall als indirekte Maße für die kognitive Bewertung dienen, sich veränderten. Zu Beginn standen Emotionen im Vordergrund, die mit der Bewertung der Situation als Bedrohung (z.B. ängstlich, besorgt) oder als Herausforderung (z.B. hoffnungsvoll, zuversichtlich) in Verbindung stehen, zu späteren Untersuchungszeitpunkten waren Emotionen im Vordergrund, die mit der Bewertung der Situation als Verlust (z.B. ärgerlich, traurig) oder Gewinn (z.B. glücklich, erleichtert) in Verbindung gebracht werden. Auch die Bewältigungsstrategien änderten sich über die Zeit. Zunächst wurden mehr problemzentrierte Strategien eingesetzt, im weiteren Verlauf aber dann mehr emotionszentrierte Strategien. Allerdings nutzten über alle Messzeitpunkte hinweg 94 Prozent der Befragten beide Formen der Bewältigung (vgl. Folkman & Lazarus, 1980). Neben diesen Veränderungen über die Zeit konnten auch interindividuelle Unterschiede festgestellt werden. Wichtige Prädiktoren waren hier, wie schwierig das Examen angesehen wurde und wie das

Ausmaß an wahrgenommener Kontrolle ausfiel. Die Befunde deuten darauf hin, dass Stress ein Prozess ist, der sich im Zeitverlauf ändert und abhängig ist von den Bewertungen, die eine Person einnimmt und den Bewältigungsstrategien, die sie einsetzt. Dennoch kann auch diese Studie aus der Berkeley-Gruppe nur als indirekter Beleg für die transaktionale Stresstheorie angesehen werden, da einerseits die Bewertungsprozesse nicht direkt erfasst, sondern nur indirekt über Emotionen erhoben wurden (vgl. Monroe & Kelley, 1997), zum anderen wurden auch hier die angenommenen mediierenden Prozesse nicht analysiert. Ein weiteres Problem ist, dass die Beziehungen zwischen Bewältigungsstrategien und Emotionen nicht über die Zeit untersucht wurden, da es möglich ist, dass bestimmte Strategien zu bestimmten Emotionen zu einem späteren Zeitpunkt führen können (vgl. Lazarus, 1991a, 1999).

Eine Replikation der Studie von Folkman und Lazarus (1985) durch Drumheller, Eike und Scherer (1991) zeigt zusätzlich, dass auch der generelle Stresslevel von Personen eine Rolle spielt und auch Ressourcen (in diesem Fall die Leistung der Studierenden in vorherigen Examina) einen Einfluss auf den Stressprozess haben. Studierende mit einem geringeren Stresslevel zeigen einen niedrigeren Bewältigungsaufwand. Bessere Leistungen in vorherigen Examina gehen mit weniger bedrohlichen Bewertungen einher. Auch in dieser Studie wurde die kognitive Bewertung nur indirekt über Emotionen erfasst, mediierende Prozesse wurden nicht analysiert.

Auch Carver und Scheier (1994) haben die Studie von Folkman und Lazarus (1985) repliziert und erweitert. Ihre Erweiterung bestand darin, dass sie die kausalen Beziehungen zwischen Bewältigungsverhalten und Emotionen untersuchen wollten. Eine weitere zusätzliche Fragestellung bezog sich auf die Unterscheidung zwischen Bewältigungsstilen und situativem Bewältigungsverhalten. Dies ist insofern relevant, da Bewältigungsstile einen Einfluss auf situatives Bewältigungsverhalten haben können. Das Muster der Emotionen über die drei Examensphasen hinweg repliziert die Ergebnisse von Folkman und Lazarus (1985). Ohne auf Details einzugehen, erweitern die Ergebnisse der Studie von Carver und Scheier (1994) die Befunde von Folkman und Lazarus insofern, dass auch kausale Effekte von bestimmten Bewältigungsstilen auf Emotionen gefunden werden konnten. Von größerer Relevanz sind die Ergebnisse zum Zusammenhang von Bewältigungsstilen und situativem Bewältigungsverhalten. Carver und Scheier finden einen moderaten Zusammenhang, allerdings haben Bewältigungsstile in ihrer Untersuchung keinen prädiktiven Wert für spezifische Emotionen. Weiterhin finden die Autoren, dass Vertrauen („confidence“) im Sinne einer personalen Ressource ein Prädiktor für positivere Emotionen (Emotionen im Zusammenhang mit Herausforderung und Gewinn) ist (s.a. Carver & Scheier, 1990; Scheier & Carver, 1992).

Eine Untersuchung von Bolger (1990) mit dem Examen als Stressor befasste sich mit der Rolle von Neurotizismus im Stressprozess. Ihn interessierte, ob Neurotizismus einen Einfluss auf das Stressausmaß hat und ob das Bewältigungsverhalten den Einfluss von Neurotizismus mediiert. Die zentralen abhängigen Variablen waren die Ängstlichkeit und die Leistung der Studierenden (Note im Examen). Die Ergebnisse zeigen, dass höhere Neurotizismuswerte mit höheren Ängstlichkeitswerten zu einem späteren Messzeitpunkt einhergehen und dass dieser Effekt durch emotionszentrierte Bewältigungsstrategien (z.B. Wunschdenken) mediiert wird. Auch die Studie von Bolger kann nur als indirekter Test für die transaktionale Stresstheorie herangezogen werden, da der zentrale Modellbaustein, die kognitiven Bewertungen, nicht erhoben wurden. Allerdings zeigen die Ergebnisse, dass personale Ressourcen bzw. in diesem Fall stabile Persönlichkeitsmerkmale über bestimmte Bewältigungsstrategien einen Effekt auf das Stressausmaß haben (s. dagegen McCrae & Costa, 1986). Anzunehmen ist weiter, dass Neurotizismus sich zunächst auf die kognitive Bewertung (Ereignis- und Ressourcenbewertung) auswirkt und es hierüber dann zur Auswahl eher emotionszentrierter bzw. abwehrender Bewältigungsstrategien kommt (Bolger & Zuckerman, 1995; Gunthert, Cohen & Armeli, 1999).

Den Stressprozess im Rahmen einer Unternehmensfusion untersuchten Fugate, Kinicki und Scheck (2002) und orientierten sich dabei an den Studien von Folkman und Lazarus (1985) und Carver und Scheier (1994). Zwar nutzten die Autoren nicht wie in den oben aufgeführten Untersuchungen ein Examen als Stressor, sondern eine Fusion zweier Unternehmen, die in verschiedene Phasen eingeteilt werden konnte und eine Vorher- und Nachhermessung (vor und nach Eintritt der Fusion) ermöglichte. Über vier Messzeitpunkte hinweg mit einem Abstand von jeweils ca. drei Monaten erhoben sie an einer Stichprobe von ArbeitnehmerInnen Angaben zu personalen Ressourcen (Ausmaß an sozialer Unterstützung, wahrgenommene Kontrolle), zur kognitiven Bewertung, zu Emotionen und zum Bewältigungsverhalten. Über den Verlauf der Fusion konnten bedeutsame Veränderungen in der Bewertung der Fusion (im Laufe der Fusion wurde die Bewertung negativer im Sinne von mehr Bedrohung und mehr Verlust), im Ausmaß an sozialer Unterstützung, in der Kontrollbewertung und im Bewältigungsverhalten festgestellt werden, die Emotionen änderten sich dagegen nicht über den Untersuchungszeitraum. Die Autoren sehen ihre Ergebnisse als Unterstützung der Annahmen der transaktionalen Stresstheorie an. Allerdings analysierten sie nicht die Zusammenhänge der Konstrukte im Sinne der Theorie über die Zeit hinweg. So werden zwar interindividuell gleich verlaufende Veränderungen über die Zeit hinweg festgestellt, wodurch diese Änderungen zustande kommen und ob es in Abhängigkeit unterschiedlicher Bewertungen und Bewälti-

gung auch interindividuelle Unterschiede in diesen Veränderungen gibt, wurde nicht untersucht.

Die hier vorgestellten Studien haben den Vorteil, über die Zeit hinweg die Anpassung an einen für alle untersuchten Personen gleichen Stressor zu analysieren. Darüber hinaus können intraindividuelle Veränderungen in bestimmten Modellvariablen wie der kognitiven Bewertung und dem Bewältigungsverhalten beobachtet oder interindividuelle Unterschiede in diesen Variablen und dessen Vorhersagekraft für Anpassungsindizes untersucht werden. Somit haben diese Untersuchungen einen wichtigen Beitrag zur Überprüfung der transaktionalen Stresstheorie geleistet. Sie zeigen, dass es Schwankungen über die Zeit in den Prozessvariablen gibt und dass diese Schwankungen Konsequenzen haben. Aber keine der aufgeführten Untersuchungen berücksichtigt in einem Analyseschritt beide Varianzquellen, intra- und interindividuell, gleichzeitig. Auch die funktionalen Relationen zwischen den verschiedenen Modellkonstrukten wurden nicht berücksichtigt.

4.2.4 Experimentelle Untersuchungen

Laborexperimentelle Untersuchungen haben den Vorteil, kontrolliert kausale Beziehungen zu prüfen und sind damit auch für eine Überprüfung der transaktionalen Stresstheorie und ihrer Annahmen von eminenter Wichtigkeit. Ein Problem von Laborexperimenten ist ihre geringe externe Validität. Die Gültigkeit der experimentell gewonnenen Ergebnisse ist oft schwer auf den Alltag übertragbar, die situativen Arrangements sind meist künstlich und treten in dieser Form nicht unter natürlichen Gegebenheiten auf. Gleichzeitig sind laborexperimentelle Untersuchungen zur Prüfung der transaktionalen Stresstheorie aufwendig. Der hohe Aufwand wird auch darin deutlich, dass es letztlich nur wenige experimentelle Untersuchungen gibt, die dann auch nur einen Ausschnitt der Lazarus'schen Theorie überprüft haben.

Jerusalem (1990, 1991) hat sich in einer aufwendigen experimentellen Untersuchung intensiv mit den kognitiven Bewertungsprozessen und deren zeitliche Entwicklung beschäftigt (s.a. die Experimente von Tomaka, Blascovich und Kollegen zu Bedrohungs- und Herausforderungsbewertung, z.B. Tomaka, Blascovich, Kelsey & Leitten, 1993; für einen Überblick Blascovich & Tomaka, 1996). Die VersuchsteilnehmerInnen sollten über neun Messzeitpunkte hinweg Leistungsaufgaben bearbeiten, nach jeder Leistungsaufgabe erhielten sie eine (fiktive) Rückmeldung über ihre Leistung (jeweils im Vergleich zu anderen TeilnehmerInnen), wobei die Stichprobe in eine Erfolgsgruppe (überdurchschnittliche Leistung) und eine Misserfolgsgruppe (unterdurchschnittliche Leistung) unterteilt wurde. Nach der Leistungsrückmeldung gaben die ProbandInnen jeweils ihre Ereignis- (Herausforderung, Bedrohung,

Verlust) und Ressourceneinschätzung (situative Leistungserwartungen und Fähigkeitsattributionen) an. Es wurden auch personale Ressourcen mit berücksichtigt (allgemeine Selbstwirksamkeitserwartungen, intelligenzspezifische Wirksamkeitsüberzeugungen, Leistungsängstlichkeit). Die Ergebnisse entsprechen in Großteilen den theoretischen Annahmen. Sowohl die Situation (Erfolg vs. Misserfolg) als auch die Person (Kompetenzerwartungen, Leistungsängstlichkeit) haben einen Einfluss auf die Ereignis- und Ressourceneinschätzungen. Misserfolgrückmeldungen gehen mit stärkeren Bedrohungs- und Verlusteinschätzungen einher als Erfolgsmeldungen, der Unterschied vergrößerte sich im experimentellen Verlauf. Ein noch stärkerer Effekt findet sich für die situationsspezifischen Fähigkeitseinschätzungen. Bei Erfolg sind die Leistungserwartungen hoch, bei Misserfolg deutlich niedriger. Auf Seiten der personalen Einflüsse zeigen sich ebenfalls signifikante Effekte: Hohe Kompetenzeinschätzungen führen zu mehr Herausforderungseinschätzungen und weniger Bedrohungs- und Verlusteinschätzungen, Leistungsängstlichkeit dagegen erzeugt höhere Bedrohungs- und Verlusteinschätzungen als auch niedrigere situationsspezifische Kompetenzerwartungen. Darüber hinaus ergeben sich bedeutsame Interaktionen. Personen mit schwachen Ressourcen reagieren besonders stark auf Misserfolgrückmeldungen, starke Ressourcen, wie besonders hohe allgemeine Selbstwirksamkeitserwartungen wirken dagegen als Puffer in solchen Situationen. Insgesamt ist diese aufwendige experimentelle Untersuchung ein starker Beleg für die Bedeutsamkeit von sowohl personalen Ressourcen als auch situationsbezogenen kognitiven Bewertungen (s. dagegen Steptoe & Vögele, 1986). Allerdings wird der gesamte Stressprozess nicht abgebildet. Die ProbandInnen haben keinen Einfluss (im Sinne des Bewältigungsverhaltens) auf den Verlauf des Experiments, die dynamische Person-Umwelt-Relation konnte somit nicht analysiert werden.

An dem komplexen experimentellen Design, das der Untersuchung von Jerusalem zugrunde liegt, wird der hohe Aufwand eines experimentellen Zugangs zur Überprüfung der transaktionalen Stresstheorie deutlich. Gleichzeitig zeigt sich auch, dass diesem Zugang Grenzen gesetzt sind. Die systematische Kontrolle von wichtigen Modellkomponenten lässt es nicht zu, den Versuchspersonen eine aktive Einflussnahme auf die experimentelle Situation einzuräumen. Würde man den Versuchspersonen einen Einfluss gewähren, würde man die Vorteile des experimentellen Zugangs aufgeben. Dennoch ist die Arbeit von Jerusalem von großer Bedeutung, kann sie doch Annahmen zur Rolle von situativen Bewertungen und deren Konsequenzen experimentell bestätigen.

4.2.5 Pfad- und Strukturgleichungsanalysen

Eine Reihe von Untersuchungen bedient sich bei der Auswertung sowohl von Quer-¹ als auch Längsschnittdaten, so genannten Pfad- und Strukturgleichungsanalysen². Diese statistischen Methoden ermöglichen es, komplexe Wirkungszusammenhänge zwischen verschiedenen Konstrukten auf ihre empirische Bewährung hin zu überprüfen (Ulman & Bentler, 2003; s. ausführlich Kap. 6.2.1). Im Folgenden werden exemplarisch Studien vorgestellt, wobei zunächst auf eine Reihe von Querschnittuntersuchungen eingegangen wird, um schließlich relevante Längsschnittuntersuchungen darzustellen.

Querschnittuntersuchungen

In einer Querschnittuntersuchung mit australischen ArbeitnehmerInnen konnten Terry, Tonge und Callan (1995) mit Strukturgleichungsanalysen Annahmen der transaktionalen Stresstheorie bestätigen. Es wurden Angaben zu personalen Ressourcen (Selbstwertschätzung, Neurotizismus, generalisierte Kontrollüberzeugungen, Ausmaß an sozialer Unterstützung), zu situativen Bewertungen (Stressausmaß, situationale Kontrollüberzeugungen, situationsspezifische Selbstwirksamkeitserwartungen) und zum Bewältigungsverhalten erhoben, zusätzlich wurden aber als Ergebnisvariablen noch die Arbeitszufriedenheit und das psychische Wohlbefinden erfasst. Ein modifiziertes Strukturgleichungsmodell belegt die Struktur der transaktionalen Stresstheorie: Die situativen Bewertungen und das Bewältigungsverhalten mediierten einen Teil der Effekte der personalen Ressourcen auf die Ergebnisvariablen. Ebenso wird ein Teil der Effekte der situativen Bewertungen auf die Ergebnisvariablen über das Bewältigungsverhalten mediiert. Um eine befriedigende Anpassung des Modells an die Daten zu erhalten, mussten aber auch direkte Pfade von den personalen Ressourcen auf das psychische Wohlbefinden und die Arbeitszufriedenheit zugelassen werden.

Scheck und Kinicki (2000) überprüften die transaktionale Stresstheorie im Kontext der Übernahme eines Unternehmens. Die Stichproben setzten sich aus ArbeitnehmerInnen der akquirierten Organisation zusammen. Erhoben wurden im Querschnitt Angaben zu personalen Ressourcen (Ausmaß an sozialer Unterstützung, wahrgenommene Kontrolle und Selbstwirksamkeitserwartungen), situationale Bedingungen (Stabilität des Arbeitsplatzes, Arbeitsplatzatmosphäre, Arbeitsplatzsicherheit), Ereigniseinschätzung (Bedrohung, Verlust, Herausforde-

¹ Querschnittdaten erlauben keine kausale Interpretation, Pfad- und Strukturgleichungsanalysen liefern lediglich Hinweise über die Struktur der Daten. Kausale Beziehungen können nur mit Längsschnitt- und experimentellen Untersuchungen geprüft werden.

² Mit Pfadanalysen kann lediglich die Struktur zwischen verschiedenen Variablen modelliert werden, Strukturgleichungsanalysen ermöglichen zusätzlich die Berücksichtigung des Messmodells der Variablen (Tabachnik & Fidell, 2001; s. Kap. 6.2.1)

rung), Emotionen (Ärger, Traurigkeit) und Bewältigungsverhalten. Strukturgleichungsanalysen zeigen, dass die Struktur der transaktionalen Stresstheorie bestätigt werden kann, wobei kleinere Modifikationen vorgenommen werden mussten: Soziale Unterstützung wirkt direkt auf das Bewältigungsverhalten und nicht vermittelt über die primäre Bewertung. Die primäre Bewertung hat direkte Effekte auf problemzentrierte Bewältigungsstrategien, die nicht über Emotionen vermittelt werden (zur Rolle von Emotionen s. Lazarus, 1991a, 1999). Allerdings wurden auch nur wenige spezifische Emotionen erhoben, Ergebnisvariablen fehlen ganz. Dennoch bestätigt sich die Struktur der Theorie in dieser Studie.

In einer weiteren Querschnittuntersuchung ist Chang (2000) mit einer Stichprobe von Studierenden am Tag eines Examens der Frage nachgegangen, ob die Effekte von Unsicherheit im Sinne eines stabilen Persönlichkeitsmerkmals (vgl. Weary & Edwards, 1994) auf depressive Symptome durch kognitive Bewertungsprozesse und Bewältigungsverhalten im Sinne der Lazarus'schen Theorie mediiert werden. Pfadanalysen zeigen, dass ein Teil der Effekte tatsächlich durch die beiden zentralen Bausteine der transaktionalen Stresstheorie vermittelt werden. Dieses Ergebnis stützt gleichzeitig die Struktur der transaktionalen Stresstheorie. Dunkley, Zuroff und Blankstein (2003) ermitteln ähnliche Ergebnisse für das Persönlichkeitskonstrukt „selbstkritischer Perfektionismus“, nutzen aber statt Pfadanalysen Strukturgleichungsanalysen. In beiden Studien wird allerdings nicht weiter untersucht, ob die kognitive Bewertung dem Bewältigungsverhalten vorausgeht, so wie Lazarus dies annimmt.

Lengua und Long (2002) konnten in einer Studie auch für 8-12jährige Kinder die Struktur der transaktionalen Stresstheorie bestätigen, sie verwendeten als Auswertungsverfahren allerdings nur Pfadanalysen. Die Effekte von negativer Emotionalität auf internalisierte (depressive Symptome) und externalisierte (Aggression) Probleme werden für diese Stichprobe über die Ereignisseinschätzung mediiert. Die Ereignisseinschätzung wiederum hat signifikante Effekte auf das Bewältigungsverhalten, welches wiederum der proximalste Prädiktor für die Probleme der Kinder ist.

Längsschnittuntersuchungen

An einer Stichprobe von ostdeutschen MigrantInnen, die im Zeitraum von August 1989 bis Februar 1990 nach Westdeutschland emigrierten, untersuchte Jerusalem (1993) mit Strukturgleichungsanalysen die transaktionale Stresstheorie mit allen Modellkomponenten. Es wurden zwei Messzeitpunkte realisiert, zum ersten Messzeitpunkt wurden personale Ressourcen und situationale Einschränkungen (allgemeine Selbstwirksamkeitserwartungen, dispositiver Optimismus, Hilflosigkeit, Arbeitslosigkeit und Wohnverhältnisse) erhoben, zum

zweiten Messzeitpunkt wurde die Ereignisseinschätzung (Bedrohungs- und Verlustbewertung), emotionales Bewältigungsverhalten und der subjektive Gesundheitszustand erfasst. Die Analysen ergaben, dass die Struktur des transaktionalen Stressmodells eine gute Anpassung an die empirischen Daten aufweist: Personale Ressourcen beeinflussen die kognitive Bewertung, die wiederum das Bewältigungsverhalten, welches schließlich den subjektiven Gesundheitszustand bedingt. Probleme der Analysen liegen darin, dass die Autokorrelationen der verschiedenen Konstrukte nicht berücksichtigt wurden und nur interindividuelle Unterschiede überprüft wurden, nicht aber intraindividuelle Veränderungen abgebildet werden, was dem prozessualen Charakter der Theorie mehr entspräche.

Hudek-Knežević und Kardum (2000) überprüften an einer Stichprobe von Frauen die Struktur der Stresstheorie. Daten zum subjektiven Ausmaß an sozialer Unterstützung, Bewältigungsstilen und situativem Bewältigungsverhalten, zur kognitiven Bewertung (Bedrohungsbewertung und Kontrolleinschätzung) und zu kurzfristigen Ergebnissen (Angaben zur Zufriedenheit mit der Bewältigung und zur eingeschätzten Effektivität der Bewältigung) wurden zu zwei getrennten Zeitpunkten erhoben (Soziale Unterstützung und Bewältigungsstile zum ersten Messzeitpunkt, kognitive Bewertung, situatives Bewältigungsverhalten und kurzfristige Ergebnisse einen Monat später). Die kognitive Bewertung und das situative Bewältigungsverhalten wurden jeweils bezogen auf vier vorgegebene Situationen erfasst und für die Analysen über alle Situationen gemittelt. Insgesamt wurden mit Strukturgleichungsanalysen vier Modelle getestet (jeweils getrennt nach Bewältigungsstilen und situationalem Bewältigungsverhalten sowie Bedrohungsbewertung und Kontrolleinschätzung). Die Analysen zeigen, dass die Bedrohungsbewertung und die Kontrolleinschätzung einen Teil der Effekte von Bewältigungsstilen auf das kurzfristige Stressergebnis medieren. Dagegen hat situatives Bewältigungsverhalten nur geringe Effekte auf das kurzfristige Ergebnis. Insgesamt zeigen die Ergebnisse die Wichtigkeit der kognitiven Bewertung für kurzfristige Ergebnisse des Stressprozesses. Problem dieser Studie ist aber die Aufteilung des eigentlichen Modells in mehrere kleine Modelle und deren getrennte Auswertung, so dass eine Gesamtbeurteilung nicht möglich ist. So bleibt zum Beispiel der Zusammenhang zwischen Bewältigungsstilen und situationalem Bewältigungsverhalten ebenso offen wie das Zusammenwirken der Bedrohungsbewertung und der Kontrolleinschätzung im Sinne der Ereignis- und Ressourceneinschätzung.

In einer weiteren Längsschnittstudie untersuchten Shewchuk, Elliot, MacNair-Semands und Harkins (1999) den Einfluss der Big Five-Persönlichkeitsvariablen (Neurotizismus, Extraversion, Offenheit für Erfahrungen, Verträglichkeit und Gewissenhaftigkeit; Costa & McCrae, 1985) auf die Bedrohungsbewertung und das Bewältigungsverhalten über die Zeit.

Über drei Messzeitpunkte hinweg wurden Studierende zu ihren Bedrohungsbewertungen und zu ihrem Bewältigungsverhalten in kurz zurückliegenden stressigen Situationen befragt. Zusätzlich wurden zum ersten Messzeitpunkt die Big Five erhoben. Strukturgleichungsanalysen zeigen, dass die Persönlichkeitsfaktoren, dabei hauptsächlich Neurotizismus, einen bedeutsamen und direkten Einfluss auf das Bewältigungsverhalten zu allen drei Messzeitpunkten haben, unabhängig von der Bedrohungsbewertung. Dies weist auf die Bedeutsamkeit von stabilen Persönlichkeitseigenschaften im Stressprozess hin. Auch die moderaten Autokorrelationen (zwischen .35 und .55) für die Bedrohungsbewertung und das Bewältigungsverhalten zeigen, dass es stabile Formen der Bewertung und der Bewältigung von stressvollen Situationen gibt. Aber die Bedrohungsbewertung ist zu jedem Messzeitpunkt eine wichtige Determinante für das Bewältigungsverhalten, was auf situationsspezifische Prozesse hinweist.

Sweet, Savoie und Lemyre (1999) überprüften in zwei unabhängigen Studien zwei konkurrierende Modelle: Ein Modell (Stichprobe 1), in dem die kognitive Bewertung einer Situation das Bewältigungsverhalten vorhersagt, welches wiederum einen Einfluss auf das Stressausmaß als kurzfristige Folge des Stressprozesses hat und ein alternatives Modell (Stichprobe 2), in dem der Effekt des Bewältigungsverhaltens über die kognitive Bewertung mediiert auf das Stressausmaß wirkt. Dieses alternative Modell soll den neueren theoretischen Formulierungen von Lazarus, der kognitiv-motivationalen-emotiven Theorie (z.B. Lazarus, 1993a) entsprechen. Allerdings kann dieses Alternativmodell so nicht aus den Formulierungen Lazarus' zu seiner kognitiv-motivationalen-Emotionstheorie entnommen werden. Vielmehr haben meiner Ansicht nach die Autorinnen Struktur und Prozess verwechselt. Das ursprüngliche Modell (Stichprobe 1) entspricht der strukturellen Annahme, während das Alternativmodell die dynamischen, prozessualen Annahmen der transaktionalen Sichtweise von Lazarus widerspiegelt: In der Auseinandersetzung eines Individuums mit einer stressigen Situation wirkt das Bewältigungsverhalten auf die Bewertungsformen zurück (vgl. z.B. Lazarus, 1991a, s.a. Kap. 3). Beide Stichproben setzen sich aus Frauen zusammen, die sich einer Brustkrebs-Untersuchung unterzogen, der Stressor war die Befürchtung, an Brustkrebs erkrankt zu sein. Angaben zur kognitiven Bewertung, zum Bewältigungsverhalten und zum Stressausmaß wurden am Tag der Untersuchung sowie 20 Tage später erhoben. Zu diesem Zeitpunkt waren die Untersuchungsergebnisse den Frauen bekannt. Strukturgleichungsanalysen zeigen, dass das Modell, welches annimmt, dass die kognitive Bewertung das Bewältigungsverhalten vorhersagt, welches wiederum das Stressausmaß beeinflusst, bessere Anpassungsindizes als das Alternativmodell aufweist. Die Ergebnisse stützen somit die Struktur der transaktionalen Stresstheorie, wenn auch personale Ressourcen nicht mit erhoben wurden.

Gowan, Riordan und Gatewood (1999) können im Rahmen eines Arbeitsplatzverlusts zeigen, dass das transaktionale Stressmodell zur Erklärung des allgemeinen Wohlbefindens der Betroffenen und zur Wiederbeschäftigung herangezogen werden kann. Sie erhoben Angaben zu personalen Ressourcen, zur kognitiven Bewertung, zum Bewältigungsverhalten und Wohlbefinden und zu einem späteren Zeitpunkt Angaben zur Wiederbeschäftigung. Mittels Strukturgleichungsanalysen können sie in Teilen die Struktur der Stresstheorie bestätigen. Personale Ressourcen wirken in Teilen vermittelt über die Bewertung auf das Bewältigungsverhalten, welches wiederum Effekte auf die Ergebnisvariablen Wohlbefinden und Wiederbeschäftigung hat. Allerdings sind nicht alle erwarteten Pfade signifikant, darüber hinaus haben personale Ressourcen auch direkte Effekte auf das Bewältigungsverhalten.

Haney und Long (1995) untersuchten Sportteams über zwei Wettkampfrunden hinweg und konnten zeigen, dass die in der ersten Wettkampfrunde gezeigte Leistung einen Einfluss auf die Ressourceneinschätzung und das Bewältigungsverhalten in der zweiten Runde und hierüber vermittelt wiederum auf die Leistung hat. Dieser Befund stützt die transaktionale Stresstheorie, da kurzfristige Ergebnisse (hier die Leistung) zu einer Neubewertung führen (revidierte Ressourceneinschätzung) und über eine Änderung im Bewältigungsverhalten eine Veränderung in der Leistung bewirken.

Die Untersuchung von Florian, Mikulincer und Taubman (1995) kann ebenfalls als Beleg für Annahmen der transaktionalen Stresstheorie herangezogen werden. Die Autoren untersuchten, ob der Effekt von Zähigkeit („hardiness“), einem Persönlichkeitsmerkmal, welches als vorbeugender Faktor von Stress angesehen wird (Kobasa, Maddi & Kahn, 1982), auf das psychische Wohlbefinden über die kognitive Bewertung und das Bewältigungsverhalten mediiert wird. Pfadanalysen können im Längsschnitt die Annahme bestätigen.

Längsschnittuntersuchungen haben im Vergleich zu Querschnittuntersuchungen den Vorteil, kausale Beziehungen mit Hilfe von Pfad- und Strukturgleichungsanalysen zu modellieren (s. hierzu ausführlich Kap. 7.2.1). Somit geben die aufgeführten Längsschnittbefunde wichtige Hinweise für die Gültigkeit der transaktionalen Stresstheorie. Problem bleibt aber, dass nur interindividuelle Unterschiede in solchen Analysen dargestellt wurden, intraindividuelle Veränderungen, so wie sie Lazarus mit seiner transaktionalen Theorie erklären möchte, wurden in den oben aufgeführten Studien nicht analysiert. Ein weiteres Manko vieler Untersuchungen, die auf Strukturgleichungsanalysen zurückgreifen, stellt eine inadäquate Verwendung des Verfahrens dar. Gestützt wird diese Kritik durch eine Überblicksarbeit über den Einsatz von Strukturgleichungsanalysen in der Stressforschung allgemein (Roesch, 1999). Zwar liegt die Publikation der Arbeit bereits fünf Jahre zurück, die Ergebnisse sind aber auch heute

noch für die meisten Studien, die in diesem Kontext Strukturgleichungsanalysen anwenden, gültig. Insgesamt hat Roesch (1999) 50 Studien gesichtet, die mit nur wenigen Ausnahmen folgende Probleme in der Anwendung von Strukturgleichungsanalysen aufweisen: (a) Die Daten werden meist nicht auf Normalverteilung geprüft, was eine wichtige Voraussetzung für die Anwendung von Strukturgleichungsanalysen ist, (b) äquivalente Modelle werden nicht aufgelistet und diskutiert, (c) Modellspezifikationen werden meist nach empirischen Gesichtspunkten, nicht nach theoretischen Überlegungen vorgenommen (zu Problemen und Voraussetzungen s. ausführlich Kap. 6.2.1). All dies kann zu falschen Schlüssen führen und letztlich die Theorienprüfung und -bildung behindern.

4.2.6 Kombination intra- und interindividueller Analyseverfahren

Für die Überprüfung der transaktionalen Stresstheorie sind Analyseverfahren von unmittelbarer Bedeutung, die intra- und interindividuelle Varianz gleichzeitig berücksichtigen. Hierdurch können intraindividuelle Veränderungen über die Zeit analysiert und gleichzeitig interindividuelle Unterschiede in diesen Veränderungen berücksichtigt werden. Mit Ausnahme der Studien von Bolger und Zuckerman (1995) und Gunthert et al. (1999) haben bislang solche Analyseverfahren zur Überprüfung der transaktionalen Stresstheorie keine Verwendung erfahren.

Gunthert et al. (1999) untersuchten den Einfluss von Neurotizismus auf alle Modellkomponenten der transaktionalen Stresstheorie, bisherige Studien untersuchten jeweils nur eine bestimmte Modellkomponente wie die Häufigkeit der Konfrontation mit stressigen Ereignissen (Affleck, Tennen, Urrows & Higgins, 1994; David, Green, Martin & Suls, 1997), die kognitive Bewertung (Gallagher, 1990), das Bewältigungsverhalten (Bolger & Zuckerman, 1995; McCrae & Costa, 1986) und die Stressergebnisse (z.B. David et al., 1997). Da die Untersuchung von Gunthert et al. (1999) eine direkte Erweiterung der Studie von Bolger und Zuckerman (1995) darstellt, wird auf deren Untersuchung nicht gesondert eingegangen. Studierende sollten über einen Zeitraum von zwei Wochen für jeden Tag das stressigste Ereignis berichten. Die Auswertung erfolgte mit einer hierarchischen linearen Analyse (Mehrebenenanalyse, z.B. Bryk & Raudenbush, 1992), die sowohl intraindividuelle als auch interindividuelle Varianz berücksichtigt. Studierende mit hohen Neurotizismuswerten gaben negativere Ereignis- (Stressausmaß des Ereignisses) und Ressourceneinschätzungen (geringe Bewältigungsmöglichkeiten) an, nutzten weniger adaptive Bewältigungsstrategien und reagierten mit einem schlechteren allgemeinen Wohlbefinden. Die Ergebnisse zeigen, dass ein Teil der

negativen Rolle von Neurotizismus im Stressprozess bzw. in der Stressreaktivität auf die Ereignis- und Ressourceneinschätzung von stressigen Situationen zurückzuführen ist. Die strukturellen Zusammenhänge zwischen den Modellkomponenten wurden allerdings nicht weiter untersucht.

4.3 Stand der empirischen Überprüfung der transaktionalen Stresstheorie: Ein Fazit

Studien, die direkt oder indirekt die transaktionale Stresstheorie überprüft haben, lassen sich unter methodischen Gesichtspunkten in sechs Kategorien unterteilen. Alle Untersuchungen weisen mehr oder minder Mängel unter methodischen Gesichtspunkten auf. Auf die wichtigsten soll an dieser Stelle eingegangen werden.

Die transaktionale Stresstheorie stellt an empirische Untersuchungen eine Reihe von Anforderungen, die weiter oben bereits beschrieben wurden und folgend noch einmal kurz zusammengefasst werden sollen. Die transaktionale Stresstheorie macht Aussagen über die strukturellen Zusammenhänge zwischen den Variablen (s. Abbildung 1): Umwelt- und Personenvariablen haben einen Einfluss auf die Bewertungsprozesse (Ereignis- und Ressourceneinschätzungen), die wiederum den Umgang (Bewältigungsverhalten) mit einer stressigen Situation bestimmen. Die Art und Weise, wie eine Person mit der stressigen Situation umgeht, hat kurz- und langfristige Konsequenzen. Gleichzeitig unterliegen der transaktionalen Sichtweise von Stress prozessuale Annahmen. Der Umgang mit stressigen Situationen wirkt auf die Bewertungsprozesse zurück (Neubewertung), was schließlich auch zu Veränderungen im Umgang mit der (veränderten) stressigen Situation führt. Beides, sowohl Struktur als auch Prozess, muss bei der Prüfung der Annahmen der transaktionalen Stresstheorie berücksichtigt werden.

Bei kritischer Sichtung der Untersuchungen, die direkt oder indirekt die transaktionale Stresstheorie überprüft haben, lässt sich feststellen, dass keine der Studien, die unter methodischen Gesichtspunkten in sechs Kategorien unterteilt wurden, diese zwei grundsätzlichen Bestandteile der Theorie gleichzeitig berücksichtigt hat.

Die strukturellen Zusammenhänge der Modellkomponenten wurden von einer Reihe von Untersuchungen analysiert. Untersuchungen, die in die Kategorie „korrelative Untersuchungen“ fallen (z.B. Chang, 1998; Pakenham, 1999; Terry, 1994), sowohl Quer- als auch prospektive (Längsschnitt-)Untersuchungen, greifen bei der Datenanalyse auf einfache Regressionsanalysen zurück, mit deren Hilfe auch einfache medierende Beziehungen zwischen Variablen analysiert werden können (Baron & Kenny, 1986). So kann beispielsweise Terry (1994) zeigen, dass die kognitive Bewertung, wie in der transaktionalen Stresstheorie angenommen,

als Mediator zwischen personalen Ressourcen und Bewältigungsverhalten fungiert. Komplexere strukturelle Zusammenhänge können nicht analysiert werden. In einfachen Regressionsanalysen kann nicht die kausale Wirkkette, so wie sie in Abbildung 1 dargestellt ist, gleichzeitig analysiert werden. Einfache regressionsanalytische Verfahren erlauben somit nur eine Analyse von Ausschnitten der Stresstheorie. Pfad- und Strukturgleichungsanalysen stellen dagegen komplexe, regressionsanalytische Verfahren dar, die genau dies ermöglichen. Eine Reihe von Untersuchungen, sowohl quer- als auch längsschnittlich, hat solche Verfahren verwendet (z.B. Jerusalem, 1993; Shewchuk et al., 1999) und können auch die Struktur der Theorie mit Einschränkungen bestätigen. Doch mit Ausnahme von Sweet et al. (1999) weisen die Untersuchungen in ihrer Anwendung der Methoden Mängel auf (z.B. berücksichtigt Jerusalem (1993) in der Auswertung seiner Längsschnittdaten nicht autoregressive Zusammenhänge; für eine Übersicht über häufig gemachte Fehler bei der Anwendung von Strukturgleichungsanalysen s. Roesch, 1999). Schließlich bieten noch laborexperimentelle Untersuchungen durch die gezielte Manipulation von Bausteinen der transaktionalen Stresstheorie die Möglichkeit, zumindest die strukturellen Annahmen zu prüfen. Die Analyse von prozessualen Annahmen dagegen ist schwieriger, da den VersuchsteilnehmerInnen die aktive Einflussnahme auf das experimentelle Setting zugestanden werden müsste, wodurch allerdings eine systematische Kontrolle der Modellvariablen verloren gehen würde. Der Aufwand für experimentelle Untersuchungen ist sehr groß, was sich auch daran ablesen lässt, dass es nur wenige laborexperimentelle Studien gibt (z.B. Jerusalem, 1990, 1991; Tomaka et al., 1993), die darüber hinaus auch nur einen Ausschnitt aus der Theorie, die kognitive Bewertung, systematisch untersucht haben. Somit lässt sich festhalten, dass es eine Reihe von Untersuchungsergebnissen gibt, die als Hinweise für die Richtigkeit der strukturellen Annahmen der transaktionalen Stresstheorie interpretiert werden können. Manko der Untersuchungen ist, dass nicht intraindividuelle Veränderungen in den Modellvariablen berücksichtigt wurden.

Es gibt Untersuchungen, die die intraindividuellen Veränderungen in zentralen Modellvariablen zum Gegenstand haben (z.B. Folkman, Lazarus, Dunkel-Schetter et al., 1986; Folkman & Lazarus, 1985) und theoriekonforme Veränderungen feststellen konnten. So können Folkman und Lazarus (1985) zeigen, dass vor einem Examen bei Studierenden Bedrohungsbewertungen dominieren, nach dem Examen und damit nach dem Eintreten des Ereignisses aber Verlustbewertungen. Allerdings werden Determinanten dieser intraindividuellen Veränderungen nicht analysiert. Wenn auf interindividuelle Unterschiede eingegangen wird, dann wird die intraindividuelle Ebene wieder außer Acht gelassen. So untersuchen auch Carver und Scheier (1994) die Veränderungen in den kognitiven Bewertungen und dem Bewältigungs-

verhalten vor und nach einem Examen. Bei der Analyse der Zusammenhänge bzw. den Beziehungen zwischen den Variablen nutzen sie nur eine korrelative Auswertung, die nur interindividuelle Unterschiede (bzw. die Stabilität von interindividuellen Unterschieden über Variablen hinweg) abbilden.

Lediglich eine der angeführten Untersuchungen greift auf ein Verfahren zurück, das sowohl die interindividuelle als auch die intraindividuelle Varianz berücksichtigt. Gunthert et al. (1999) verwenden ein hierarchisches lineares Verfahren (auch Mehrebenenanalysen genannt) zur Auswertung ihrer Längsschnittdaten, die beide Varianzquellen, also interindividuell (Unterschiede zwischen Personen in bestimmten Variablen) als auch intraindividuell (Veränderungen der Werte einer Person in einer Variable über die Zeit), berücksichtigt. Mittels dieses Analyseverfahrens konnten die Autoren zeigen, dass Neurotizismus als differentiopsychologische Variable die Veränderung in anderen Variablen, z.B. subjektive Bewältigungsmöglichkeiten, vorhersagen. Dieses Ergebnis bestätigt hinsichtlich zweierlei, bereits oben aufgeführter Punkte die transaktionale Stresstheorie. Zum einen zeigen die intraindividuellen Veränderungen in Modellvariablen den Stressprozess an (der also nicht statisch ist), zum anderen gibt es in diesen Veränderungen interindividuelle Unterschiede. Dennoch bleibt auch in dieser Untersuchung ein wichtiger Punkt ungeprüft. Das Wirkungsgefüge bzw. die strukturellen Zusammenhänge können mit hierarchischen linearen Verfahren nicht analysiert werden, da sie wie normale multiple bzw. hierarchische Regressionsanalysen nur die inkrementellen Varianzanteile in einer abhängigen Variablen anzeigen, die die Prädiktoren erklären können.

Somit kann als Fazit für den Stand der Überprüfung der transaktionalen Stresstheorie von Lazarus festgehalten werden, dass es eine Reihe von empirischen Befunden gibt, die als Teilbelege der Theorie gewertet werden können. Eine umfassende empirische Überprüfung, die sowohl die strukturelle Beziehung zwischen den Modellvariablen als auch den transaktionalen Stressprozess überprüft hat, fehlt. Ein Grund hierfür sind sicherlich die hohen methodischen Anforderungen, die eine Überprüfung der Theorie mit sich bringt. Sowohl an die Form der Daten (aufwendige Längsschnittuntersuchungen) als auch an die Auswertungsmethode stellt die Theorie hohe Ansprüche. Hierin ist auch die Ursache dafür zu suchen, dass es nur sehr wenige aktuelle Untersuchungen in diesem Bereich gibt, die Mehrzahl der Untersuchungen stammt aus den 80er und 90er Jahren des letzten Jahrhunderts.

Ziel der vorliegenden Arbeit ist, diese Lücke in der empirischen Forschung zur transaktionalen Stresstheorie zu schließen. Als Kontext wird dabei das Lehramtsreferendariat gewählt, welches den Makrostressor bildet.

5 Kontext und Fragestellung der Arbeit

Im folgenden Abschnitt wird zunächst auf den Kontext der Arbeit eingegangen, das Lehramtsreferendariat (Kap. 5.1). Dabei wird die Situation beschrieben, mit der LehramtskandidatInnen konfrontiert sind. Auch relevante empirische Untersuchungen zu Stress im Lehrberuf im Allgemeinen und im Referendariat im Speziellen werden zusammengefasst. Anschließend wird die Fragestellung der Arbeit konkretisiert (Kap. 5.2).

5.1 Kontext der Arbeit: Das Lehramtsreferendariat

Das Lehramtsreferendariat wird oft als belastende Phase in der LehrerInnenausbildung gesehen. So übertitelt Schedensack (1995) einen Beitrag zu diesem Thema mit der Frage „Referendariat als persönliche Krise?“. Er berichtet von seinen Erfahrungen als Ausbilder in einem hessischen Studienseminar zum Umgang von ReferendarInnen mit Belastungen, mit denen sie in der Referendariatszeit konfrontiert werden. Er beschreibt krisenhafte Einbrüche von ReferendarInnen und kommt zu dem Ergebnis, dass das Referendariat als Übergangsphase ein krisenanfälliger Abschnitt im Laufe der persönlichen und beruflichen Entwicklung von LehrerInnen ist. Aber auch die ReferendarInnen selbst sehen das Referendariat als sehr belastend an (Böhmman, 2000). Neben dem Leistungsdruck, der dem Referendariat inhärent ist, gibt es eine Reihe von weiteren Faktoren, die zu einer starken subjektiven Belastung bei den ReferendarInnen führen können: Sie müssen an ihrer jeweiligen Ausbildungsschule erstmals eigenständigen Unterricht halten und werden somit auch mit den im Lehrberuf bereits bekannten Belastungsfaktoren wie z.B. zu große Klassen, Vor- und Nachbereitung des Unterrichts, unterschiedliche Lernvoraussetzungen der SchülerInnen, mangelnde Motivation bzw. Konzentrationsfähigkeiten der SchülerInnen konfrontiert (vgl. van Dick, 1999a). Weiterhin ist davon auszugehen, dass es im Referendariat spezifische Belastungsfaktoren gibt, mit denen sich die Referendarinnen und Referendare auseinandersetzen müssen. Dazu zählen Konflikte mit der Ausbildungsschule, Konflikte mit den Ausbildern, Lehrproben, etc. (Oesterreich, 1987; Ulich, 1996).

Um die angehenden LehrerInnen auf die Schwierigkeiten und psychischen Belastungen vorzubereiten, gibt es eine Reihe von „Ratgebern“, die solch vielsagende Titel haben wie „Der Mutmacher“ (Handke, 1997), „Fit für das Referendariat“ (Böhmman, 2001) oder „Kursbuch Referendariat“ (Daschner & Drews, 2000). Inhalt dieser Ratgeber sind meist Tipps und Ratschläge, wie ReferendarInnen mit den Belastungen im Referendariat umgehen sollen, ohne theoretische und empirische Fundierung, was allerdings die Autoren auch so explizieren.

Lehrerstress, so der wissenschaftliche Terminus für Belastung und Beanspruchung im Lehrerberuf, ist in der Forschung seit längerer Zeit Thema einer Vielzahl von empirischen Untersuchungen (Kyriacou, 2001). Dagegen gibt es nur wenige empirische Arbeiten, die sich mit Stress und dessen Folgen im Lehramtsreferendariat beschäftigen. Bevor ein kurzer Überblick über wichtige Forschungsbefunde gegeben wird, sollen zunächst die Struktur und die Inhalte des Lehramtsreferendariats beschrieben werden.

5.1.1 Struktur und inhaltlicher Aufbau des Lehramtsreferendariats

Das Lehramtsreferendariat unterteilt sich in Hessen³ in vier Ausbildungsabschnitte, die auf insgesamt 24 Monate verteilt sind. In Tabelle 1 sind die Ausbildungsabschnitte und die jeweilige Dauer aufgeführt.

Tabelle 1. Ausbildungsabschnitte und deren Dauer im Lehramtsreferendariat

Ausbildungsabschnitt	Dauer (in Monaten)
Einführungsphase	3
Differenzierungsphase	6
Intensivphase	12
Vorbereitungsphase	3

Die Intensivphase wird meist in eine Intensivphase I und Intensivphase II unterteilt, die je gleich lang andauern.

In allen Phasen finden an den Studienseminaren Seminarveranstaltungen mit erziehungs- und gesellschaftswissenschaftlichem Arbeitsschwerpunkt (z.B. Unterrichtsplanung, Sozialformen des Unterrichts, Beurteilung von Schülerleistung, etc.), mit fachdidaktischem Arbeitsschwerpunkt (Seminare zu spezifischen Unterrichtsfächern) und mit anderen ausbildungsrelevanten Arbeitsschwerpunkten (z.B. Schul- bzw. Beamtenrecht) statt. Daneben ist ein weiterer wichtiger Schwerpunkt die unterrichtspraktische Ausbildung, die sich ebenfalls über alle Ausbildungsphasen, allerdings in einem unterschiedlichen Ausmaß erstreckt. In der Einfüh-

³ Bei der Beschreibung der Struktur und der Inhalte des Lehramtsreferendariats wird nur auf Hessen Bezug genommen, die Stichprobenziehung erfolgte nur in diesem Bundesland. Da die Schulgesetzgebung Ländersache ist, gibt es zwischen den Bundesländern Unterschiede in der Ausgestaltung des Referendariats. Dennoch sind auch viele Überschneidungen festzustellen, da die ständige Konferenz der Kultusminister der Länder vereinheitlichte Empfehlungen erarbeiten, um die Kompatibilität der Länderverordnungen zu gewährleisten. Ein Vergleich der Ausbildungsstrukturen und -inhalte findet sich bei Durdel (2000).

rungsphase umfasst die unterrichtspraktische Ausbildung 10 Wochenstunden und wird in Form von Hospitationen und angeleitetem Unterricht durchgeführt. In der Differenzierungsphase müssen die angehenden LehrerInnen zum ersten Mal eigenverantwortlich Unterricht abhalten. Insgesamt umfasst in dieser Phase die unterrichtspraktische Ausbildung 12 Wochenstunden (6 bis 8 Wochenstunden eigenverantwortlicher Unterricht, 4 bis 6 Wochenstunden Hospitation oder angeleiteter Unterricht). Während der Intensivphase erhöht sich der unterrichtspraktische Teil auf 16 Wochenstunden, wobei 10 bis 12 Wochenstunden in Form von eigenverantwortlichem Unterricht und 4 bis 6 Wochenstunden in Form von Hospitation oder angeleitetem Unterricht abgeleistet werden müssen. Schließlich reduziert sich der Anteil an Unterricht in der Vorbereitungsphase, in der die ReferendarInnen sich auf das zweite Staatsexamen vorbereiten müssen, auf 10 Wochenstunden, die in Form von Hospitation oder angeleitetem Unterricht absolviert werden können oder auf Wunsch auch in Form von eigenverantwortlichem Unterricht.

Ziel des Lehramtsreferendariats ist die Befähigung der ReferendarInnen für die vorgesehene Tätigkeit in der Schule. Abgeschlossen wird das Referendariat mit dem zweiten Staatsexamen, wobei hier überprüft werden soll, ob die/der ReferendarIn das Ziel der pädagogischen Ausbildung erreicht hat und die Befähigung für das Lehramt besitzt, für welches sie/er ausgebildet wurde. Das Staatsexamen setzt sich aus einer pädagogischen Prüfungsarbeit (Hausarbeit, i.d.R. über selbst durchgeführten Unterricht), zwei Prüfungslehrproben und einer 60minütigen, mündlichen Prüfung zusammen. Die Pädagogische Prüfungsarbeit wird meist drei Monate vor Ende der Intensivphase begonnen und an deren Ende abgegeben. Ab der Mitte der Vorbereitungsphase werden dann die Prüfungslehrproben und die mündliche Prüfung abgelegt. Neben den drei Prüfungen fließt in die Abschlussbewertung auch noch die Bewertung des Ausbildungsstandes durch die/den AusbildungsleiterIn und die/den LeiterIn der Ausbildungsschule ein.

Struktur und Inhalte der Lehrerausbildung sind in den letzten Jahren, nicht zuletzt wegen der PISA-Studie (z.B. Lange, 2002), in Kritik geraten. ExpertInnenkommissionen sowohl auf Bundesebene (Terhart, 2000; Wissenschaftsrat, 2001) als auch auf Länderebene (z.B. für Hessen: Expertengruppe Lehrerbildung, 2002) haben Änderungsvorschläge für eine Umstrukturierung der LehrerInnenausbildung gemacht.

5.1.2 Belastung und Beanspruchung von ReferendarInnen

Eine eigene Literaturrecherche in den gängigen Literaturdatenbanken (Psyndex, PsycInfo, Eric, FIS Bildung) mit den Stichwörtern „Referendariat, Referendare, Belastung, Stress, trainee teachers, student teachers, stress, strain“ hat keine Fundstellen für empirische deutsch-

sprachige und nur für wenige internationale Studien ergeben. Somit kann im Folgenden nur auf die Ergebnisse und Erkenntnisse der internationalen Forschung eingegangen werden.

Foki dieser Untersuchungen sind die Identifizierung von Belastungsfaktoren, die Ermittlung des Belastungsausmaßes und die Folgen für die Lehrleistung (Aitken & Mildon, 1991; Elkerton, 1984; Gorrell, Bregman, McAllister & Lipscomb, 1985; Houtman & Bakker, 1991a; Kaunitz, Spokane, Lissitz & Strein, 1986; MacDonald, 1992; MacDonald, MacKinnon, Joyce & Gurney, 1992; Murray-Harvey et al., 2000). Übereinstimmend zeigen die Studien, dass das Referendariat bzw. die Schulpraktika in der Lehrerausbildung mit hohen Belastungen einhergehen. Belastungsfaktoren, die von Seiten der untersuchten Referendarinnen und Referendare genannt werden, sind z.B. Disziplin der SchülerInnen, Konflikte mit den AusbilderInnen und MentorInnen, Lehrproben und Unterrichtsvorbereitung (Admiraal, Korthagen & Wubbels, 2000; Kaunitz et al., 1986; MacDonald, 1992; Morton, Vesco, Williams & Awender, 1997).

Mögliche Moderator- und Mediatorvariablen im Belastungs-Beanspruchungs-Gefüge und gesundheitliche Konsequenzen wurden auch in diesen Studien bislang nur selten berücksichtigt. Sadowski und Blackwell (1985) untersuchten den Einfluss von Kontrollüberzeugungen auf das wahrgenommene Belastungsausmaß bei ReferendarInnen. Dabei zeigt sich, dass internale Kontrollüberzeugungen mit einem geringeren subjektiven Stressausmaß einhergehen (im Vergleich zu externalen Kontrollüberzeugungen) und somit einen protektiven Faktor darstellen. Chan (2002) identifizierte das Ausmaß an sozialer Unterstützung und Selbstwirksamkeitserwartungen als protektive Faktoren. Andere Untersuchungen haben sich mit Bewältigungsstrategien der ReferendarInnen beschäftigt, d.h. wie ReferendarInnen mit den Belastungsfaktoren umgehen und welche Strategien adaptiv bzw. maladaptiv sind (Admiraal et al., 2000; Houtman & Bakker, 1991b; MacDonald, 1993; Murray-Harvey et al., 2000). Allerdings kommen die hier relevanten Untersuchungen, gerade auch wegen unterschiedlichen Forschungsstrategien (quantitative Ansätze, qualitative Befragungen) zu keinem konsistenten Ergebnismuster, so dass offen bleibt, welche Bewältigungsstrategien nützlich sind. Ein weiteres Problem stellt die Übertragbarkeit von Forschungsbefunden aus dem anglo-amerikanischen Raum auf den deutschen Kontext dar, da die Lehrerausbildung sich zum Teil stark von der in Deutschland unterscheidet (Cobb, 1999). In den USA müssen Lehreranwärter beispielsweise lediglich 8-15 Wochen Schulpraktika durchführen, während in Deutschland das Referendariat über volle 2 Jahre geht und die ReferendarInnen somit über einen längeren Zeitraum mit den Belastungen, die mit dem Unterricht einhergehen, konfrontiert werden. Darüber hinaus ist die Dichte der Leistungsüberprüfungen (Lehrproben, Examensarbeit, Prüfungen) in Deutschland größer. Dennoch lässt sich zusammenfassend festhalten, dass die Lehrerausbil-

dung, gerade wenn auch unterrichtspraktische Aspekte mit hinzukommen wie im Lehramtsreferendariat in Deutschland, mit Belastungen einhergeht.

5.1.3 Belastung und Beanspruchung von LehrerInnen

Sowohl deutsche als auch internationale Untersuchungen kommen zu dem Ergebnis, dass LehrerInnen über ein hohes Ausmaß subjektiver Belastung berichten (z.B. Boyle, Borg, Falzon & Baglioni, 1995; Capel, 1991; Kramis-Aebischer, 1995; van Dick, 1999a). Die wichtigsten Belastungsfaktoren, die in den meisten Untersuchungen erhoben und von einem substantziellen Teil der Lehrer als belastend empfunden werden, sind nach van Dick (1999a, S. 140-141) zu große Klassen, mangelnde Motivation bzw. Konzentrationsfähigkeit der SchülerInnen, geringe Lernbereitschaft von SchülerInnen, Disziplinprobleme, hektisches und störungsvolles Arbeitsklima an der Schule, unterschiedliche Lernvoraussetzungen der SchülerInnen, ständige Kritik am Lehrerberuf und fehlende Anerkennung in der Öffentlichkeit, Verwaltungsarbeit, Probleme mit den Eltern, fachfremder Unterrichtseinsatz, Vor- und Nachbereitung des Unterrichts, Ärger mit Behörden und Institutionen, Probleme mit ausländischen SchülerInnen.

Es ist anzunehmen, dass zumindest in Teilen ReferendarInnen ebenfalls mit diesen Belastungsfaktoren konfrontiert werden, da sie im Referendariat zum ersten Mal selbst unterrichten.

Gesundheitliche Folgen der empfundenen starken Belastung drücken sich bei LehrerInnen sowohl in objektiven als auch subjektiven Maßen aus (für eine Übersicht s. Hillert & Schmitz, 2004). In Hessen ließ sich 1995 beispielsweise fast die Hälfte (48,6%) aller ausscheidenden Lehrer aus „Gründen der Dienstunfähigkeit“ vorzeitig pensionieren (van Dick, 1999a; s. Weber, 2004 für gesamtdeutsche Zahlen). Jehle (1997) kommt bei der Durchsicht der vorhandenen amtlichen Statistiken zu dem Ergebnis, dass in den meisten Bundesländern in den letzten 10-15 Jahren eine Zunahme der frühzeitigen Pensionierungen festzustellen ist. Dabei leiden vorzeitig in den Ruhestand gegangene LehrerInnen im Vergleich zu allen übrigen vorzeitig in Rente gegangenen angestellten Arbeitnehmern häufiger unter psychiatrischen Erkrankungen und Krankheiten der Sinnesorgane. Scheuch und Vogel (1993) können zeigen, dass LehrerInnen bei vielen Krankheitsformen ein größeres Erkrankungsrisiko haben als andere Arbeitnehmer. Internationale Studien kommen zu ähnlichen Befunden. Zum Beispiel ermitteln Fletcher und Payne (1982) bei englischen LehrerInnen im Vergleich zur Normstichprobe sehr viel höhere Depressionswerte. Bei van Dick (1999a) korreliert in einer Stichprobe von 992 LehrerInnen die wahrgenommene Belastung zu $r = .42$ mit der Wahrnehmung diver-

ser körperlicher Beschwerden. In einer Leserumfrage durch Kretschmann (1994) in der Zeitschrift „Päd Extra“ äußern 126 deutsche LehrerInnen vor allem Schlaflosigkeit, Verspannungen, Rücken- und Wirbelsäulenprobleme sowie Magen- und Darmprobleme. Rudow (1994) fasst die Befundlage insgesamt folgendermaßen zusammen (S. 37): „...lässt sich zur psychischen Gesundheit von Lehrern eindeutig feststellen: Psychische Störungen und Befindlichkeitsbeeinträchtigungen treten am häufigsten auf. Epidemiologische Studien aus verschiedensten Ländern lassen den Schluss zu, dass darüber international ein Konsens besteht. Anscheinend sind diese Störungen bzw. Beeinträchtigungen bei Lehrern im Vergleich zu anderen Populationen oder Berufsgruppen stärker ausgeprägt.“

Zur Erklärung von Belastungen im Lehrerberuf gibt es eine Vielzahl von Untersuchungen (Kyriacou, 2001). Dabei greift ein Teil der Untersuchungen auf die transaktionale Stresstheorie von Lazarus (Lazarus, 1991a; Lazarus & Folkman, 1984) zurück. Kyriacou und Sutcliffe (1978) haben dieses Modell an den Lehrerberuf adaptiert. Sie definieren Stress im schulischen Kontext als „...a response of negative affects (such as anger or depression) by a teacher usually accompanied by potentially pathogenic physiological and biochemical changes (such as increased heart rate or release of adrenocorticotrophic hormones in the bloodstream) resulting from aspects of the teacher’s job and mediated by the perception that the demands made upon the teacher constitute a threat to his self-esteem or well-being and by coping-mechanisms activated to reduce the perceived threat“ (Kyriacou & Sutcliffe, 1978, S. 2). Allerdings zeigt sich bei genauerer Betrachtung des Modells von Kyriacou und Sutcliffe, dass keine gravierenden Neuerungen im Vergleich zu dem ursprünglichen Modell von Lazarus und Mitarbeitern vorgenommen wurde, somit können beide Modelle als äquivalent bezeichnet werden.

Exemplarisch sollen folgend Untersuchungen zu Lehrerstress vorgestellt werden, die zum Teil explizit die transaktionale Stresstheorie zur Grundlage haben, zum Teil bestimmte Modellvariablen untersucht haben, ohne sich direkt auf die Theorie zu beziehen.

5.1.4 Anwendung des Stressmodells in der Lehrerstress-Forschung

Einige Studien beschäftigen sich mit dem Zusammenhang zwischen sozialer Unterstützung (als personale Ressource) und Stress (z.B. Burke & Greenglass, 1993; Burke, Greenglass & Schwarzer, 1996; Greenglass, Fiksenbaum & Burke, 1994; Pierce & Molloy, 1990; van Dick, Wagner, Petzel, Lenke & Sommer, 1999). Dabei zeigt sich, dass soziale Unterstützung sowohl direkte als auch indirekte Effekte auf das Wohlbefinden haben kann. Andere Studien fokussierten auf die Kompetenzerwartungen der LehrerInnen und deren Einfluss auf Stress (z.B. Bhagat & Allie, 1989; Punch & Tuettemann, 1990; van Dick, Wagner & Petzel, 1999). Hierbei zeigt sich, dass sich hohe Kompetenzerwartungen negativ auf das Ausmaß des wahr-

genommenen Stress auswirken. Je höher die Kompetenzerwartungen einer/s LehrerIn/s, desto geringer das Stressausmaß. Auch Selbstwirksamkeitserwartungen und ihr Einfluss auf Belastung und Stress wurden in einer Reihe von Studien untersucht (z.B. Schmitz & Schwarzer, 2000). Übereinstimmend kommen diese Studien zu dem Ergebnis, dass eine hohe Selbstwirksamkeitsüberzeugung mit einem geringeren Ausmaß an Stress und besserer Gesundheit einhergeht.

Es gibt eine Reihe von Versuchen, mit Hilfe von Strukturgleichungsmodellen Hinweise auf die kausalen Zusammenhänge zwischen den oben vorgestellten Konzepten zu finden (Boyle, Borg, Falzon & Baglioni, 1995; Starnamann & Miller, 1992; van Dick & Wagner, 2001). Van Dick und Wagner (2001) überprüften in ihrer Studie an zwei deutschen LehrerInnenstichproben in Anlehnung an das Stressmodell von Kyriacou und Sutcliffe (1978) die Zusammenhänge mit Hilfe von Pfadanalysen. Dazu erhoben sie Daten zum Stressausmaß (Arbeitsbelastung und Mobbing), Beanspruchungen (körperliche Beschwerden und Burnout) und einige Moderatorvariablen (Selbstwirksamkeitserwartungen, Soziale Unterstützung und Stressbewältigungsstrategien). Die Ergebnisse zeigen, dass ein höheres Stressausmaß zu mehr körperlichen Symptomen führt, während ein höheres Ausmaß an sozialer Unterstützung das Stressausmaß reduziert. Soziale Unterstützung und Selbstwirksamkeitserwartungen wirken moderierend auf die Beziehung zwischen dem Stressausmaß und Burnout und körperlichen Symptomen als Stressergebnis, problemzentrierte Bewältigungsstrategien wirken reduzierend auf Burnout.

Zusammenfassend lässt sich sagen, dass die Lehrerstress-Forschung in Teilen auf die transaktionale Stresstheorie zur Erklärung von Lehrerstress zurückgegriffen hat und sich dabei die Theorie als fruchtbar erwiesen hat. Eine systematische Modelltestung wurde in diesem Bereich aber nicht vorgenommen.

5.1.5 Das Lehramtsreferendariat als Stressor

Das Referendariat dient in der vorliegenden Arbeit als Stressor, dem unausweichlich alle angehenden LehrerInnen ausgesetzt sind. Vorteil für die Untersuchung der transaktionalen Stresstheorie ist, dass der Stressor mehr oder weniger für alle ReferendarInnen gleich ist (vgl. die Examensstudien, z.B. Folkman & Lazarus, 1985). Wenn es auch in Deutschland keine systematischen Untersuchungen zum Belastungsausmaß von ReferendarInnen gibt, so machen die Beschreibungen des Referendariats als krisenhafte Phase (Schedensack, 1995) und die wenige internationale Forschung deutlich, dass die unterrichtspraktische Ausbildung von LehramtskandidatInnen (neben vielen weiteren Belastungsfaktoren) eine belastende Situ-

ation über einen längeren Zeitraum darstellt, mit der die betroffenen Personen umgehen müssen. Somit erscheint das Referendariat für die Überprüfung der transaktionalen Stresstheorie gut geeignet. Ein Nebenprodukt der vorliegenden Arbeit, welches aber als nicht weniger wichtig als das eigentliche Ziel eingeschätzt werden sollte, ist die erstmalige empirische (und theoriegeleitete) Untersuchung von Belastungsfaktoren im Referendariat und psychologischen Variablen, die einen Einfluss auf das Belastungsausmaß haben. Somit können die Ergebnisse dieser Arbeit auch direkte Implikationen für Präventivmaßnahmen haben. In der abschließenden Diskussion wird hierauf näher eingegangen werden (s. Kap. 8.3).

5.2 Fragestellung der Arbeit

In Abbildung 1 wurde die Struktur der Theorie schematisch dargestellt. Situationale und personale Variablen haben einen Einfluss auf die Ereignis- und die Ressourceneinschätzung, die wiederum das Bewältigungsverhalten bedingen. Schließlich hat die Art und Weise, wie eine Person mit einer stressigen Situation umgeht, kurzfristige und langfristige Konsequenzen. Zentrales Ziel der vorliegenden Arbeit ist die Überprüfung der Struktur der transaktionalen Stresstheorie und der prozessualen Beziehung zwischen den Modellkomponenten, die in Abbildung 1 schematisch dargestellt sind. Darüber hinaus werden auch gezielte Annahmen zu den Verknüpfungen der einzelnen Modellvariablen basierend auf empirischen Befunden formuliert. Die Schwachstellen bisheriger empirischer Untersuchungen, die in Kapitel 4 dargestellt wurden, sollen mit der vorliegenden Arbeit geschlossen werden. Hierzu wird auf statistische Auswertungsmethoden zurückgegriffen, die sowohl die Struktur als auch den Prozess analysieren können (s. Kap. 6.2.1 und 7.2.1).

Als Untersuchungskontext wurde das Lehramtsreferendariat gewählt, welches durch seine Dauer (24 Monate) als langfristiger Stressor gelten kann und somit mittelfristig Auswirkungen auf das psychische Wohlbefinden und somatische Symptome haben sollte. Aus diesem Grund sind die zentralen abhängigen Variablen der vorliegenden Arbeit das psychische Wohlbefinden und körperliche Beschwerden als eher mittel- bis langfristige Konsequenzen des Stressprozesses. Kurzfristige, d.h. physiologische und kurze emotionale Ereignisse, werden nicht untersucht. Da nur auf Selbstberichte als Datenquellen zurückgegriffen wird, werden die Auswirkungen auf das Sozialverhalten ebenfalls nicht näher untersucht, da hier externe Quellen oder Verhaltensbeobachtungen besser geeignete Verfahren darstellen. Auch für die Prozessvariablen der Theorie, also die kognitive Bewertung und das Bewältigungsverhalten, hat ein lang andauernder Stressor Konsequenzen. Von Interesse ist der generelle Bewertungs- und Bewältigungsprozess, ein längerer Zeitabschnitt wird untersucht. Kurze, auf

ein Mikroereignis bezogene Stressepisoden sollen nicht untersucht werden. Auf Seiten der situationalen und personalen Ressourcen werden als wichtige Variablen das Ausmaß an sozialer Unterstützung (Holohan & Moos, 1987; Lazarus & DeLongis, 1983; Moos & Billings, 1982; Schröder & Schwarzer, 1997; Taylor & Aspinwall, 1996; Valentiner, Holahan & Moos, 1994) und Neurotizismus (MacCrae & Costa, 1986; Vollrath & Torgersen, 2000) berücksichtigt.

Das Ausmaß an sozialer Unterstützung soll sich ebenfalls auf die kognitive Bewertung auswirken (DeLongis, Folkman & Lazarus, 1988), auch wenn Untersuchungen bislang (z.B. Terry et al., 1995) diesen Effekt nicht zeigen konnten. Es wird angenommen, dass ein hohes Ausmaß an sozialer Unterstützung sich positiv auf die kognitive Bewertung auswirkt.

Nach Bolger und Zuckerman (1995) weisen Personen mit hohen Neurotizismuswerten eine erhöhte Stressreaktivität auf. In ihrer Untersuchung können sie zeigen, dass Neurotizismus mit einer größeren Stressreaktivität einhergeht und einen Einfluss auf das Bewältigungsverhalten hat. Ihre Vermutung ist, dass die unterschiedliche Stressreaktivität durch den Einfluss von Neurotizismus auf die kognitive Bewertung verursacht wird. Gunthert et al. (1999) können dies bestätigen: Personen mit hohen Neurotizismuswerten zeigen eine negativere Ereignis- und Ressourceneinschätzung als Personen mit niedrigen Neurotizismuswerten. Auch in der vorliegenden Arbeit wird dieser negative Effekt von Neurotizismus auf die kognitive Bewertung vermutet.

Neurotizismus (oder auch negative Affektivität) wird von einigen Autoren als eine Persönlichkeitsvariable angesehen, die die Zusammenhänge zwischen verschiedenen Stressvariablen künstlich erhöht und aus diesem Grund in Untersuchungen zu Stress kontrolliert werden sollte (Brief, Burke, George, Robinson & Webster, 1988; Burke, Brief & George, 1993; McCrae, 1990; Watson & Pennebaker, 1989; s. dagegen Spector, Zapf, Chen & Frese, 2000). So finden u.a. Studien einen verzerrenden Effekt von Neurotizismus auf das wahrgenommene Ausmaß an sozialer Unterstützung (z.B. Bolger & Eckenrode, 1991; Kahn & Hessling, 2001): Je höher die Ausprägung in Neurotizismus, desto geringer fällt das subjektive Ausmaß an sozialer Unterstützung aus. Allerdings können Kahn, Hessling und Russel (2003) zeigen, dass auch bei Kontrolle von negativer Affektivität bedeutsame direkte Effekte von sozialer Unterstützung auf das Wohlbefinden bestehen bleiben. Es wird aber auch für diese Untersuchung ein negativer Einfluss von Neurotizismus auf das Ausmaß an sozialer Unterstützung angenommen.

Von der kognitiven Bewertung einer Situation hängt nach der transaktionalen Stresstheorie die Auswahl bestimmter Bewältigungsstrategien ab. Die Bewertung eines Ereignisses als

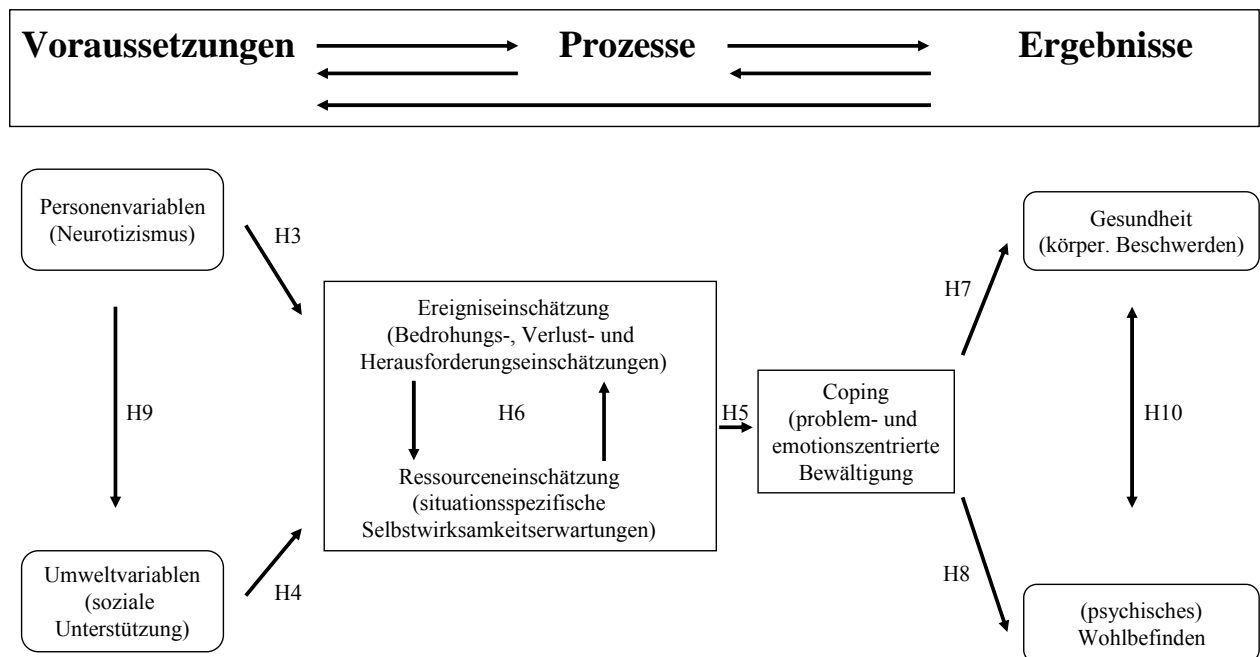
Bedrohung oder Verlust geht sowohl mit einer verstärkten Nutzung von problem- als auch emotionszentrierten Bewältigungsstrategien einher, während die Bewertung einer Situation als Herausforderung insgesamt weniger intensive Bewältigungsbemühungen evoziert und vorwiegend auf problemzentrierte Strategien zurückgegriffen wird (Carver & Scheier, 1994; Folkman & Lazarus, 1985). Andere Studien zeigen, dass auf Bedrohungs- und Verlustbewertungen die Nutzung von emotionszentrierten Strategien folgt (Bjorck & Cohen, 1993; McCrae, 1984, 1989). Basierend auf diesen Befunden wird angenommen, dass eine Bedrohungs- und/oder Verlustbewertung im Zusammenhang mit dem Referendariat mit verstärkten Bewältigungsbemühungen einhergeht, während der Zusammenhang bei einer Herausforderungsbewertung geringer ausfallen sollte. Die Ressourceneinschätzung, dabei vor allem situationsspezifische Selbstwirksamkeitserwartungen zeigen durchgängig in Untersuchungen einen positiven Zusammenhang mit problemzentrierten und einen negativen mit emotionszentrierten Bewältigungsstrategien auf (Ashford, 1988; Terry, 1994; Terry et al., 1995). Allgemeine Selbstwirksamkeitserwartungen als personale Ressourcen haben sich generell als funktional im Stressprozess erwiesen (Schröder & Schwarzer, 1997; Schwarzer, 2000). In dieser Arbeit werden sie in ihrer situationsspezifischen Form als Indikator für die Ressourceneinschätzung gesehen (vgl. Terry, 1994; s. dazu auch die Ausführung von Folkman, Lazarus, Dunkel-Schetter et al., 1986 zu „macro level“ und „micro level measures“). Es wird auch für den vorliegenden Untersuchungskontext erwartet, dass situationsspezifische Selbstwirksamkeitserwartungen positiv mit problemzentrierten und negativ mit emotionszentrierten Bewältigungsformen zusammenhängen. Darüber hinaus wird erwartet, dass die Ereignis- und Ressourceneinschätzungen miteinander zusammenhängen: Höhere Selbstwirksamkeitserwartungen sollten die Effekte der Herausforderungsbewertung auf das Bewältigungsverhalten verstärken und die Effekte von Bedrohungs- und Verlustbewertungen abschwächen (Lazarus & Folkman, 1984), wenn auch bislang solche Interaktionen selten untersucht wurden. Meist wurden nur unabhängig voneinander die Effekte der Ereignis- und Ressourceneinschätzung analysiert (z.B. Folkman, Lazarus, Dunkel-Schetter et al., 1986; Terry, 1991, 1994).

Schließlich sollen sich problemzentrierte und emotionszentrierte Bewältigungsstrategien auf das psychische Wohlbefinden und körperliche Beschwerden auswirken. Eine Reihe von Studien zeigen, dass problemzentrierte Formen von Bewältigung mit positiveren Ergebnissen einhergehen, während emotionszentrierte Formen eher mit negativen Ergebnissen zusammenhängen (Aldwin & Revenson, 1987; Aspinwall & Taylor, 1992; Folkman, Lazarus, Gruen & DeLongis, 1986). Allerdings gibt es auch Evidenz dafür, dass problemzentrierte Strategien maladaptiv sein können (Bolger, 1990; Carver & Scheier, 1994; Mattlin, Wethington &

Kessler, 1990; für einen Überblick s. Folkman & Moskowitz, 2004). Auf Grund der heterogenen Befundlage werden keine gerichteten Annahmen formuliert.

In Abbildung 2 sind in Anlehnung an Abbildung 1 alle spezifischen Untersuchungshypothesen zusammengefasst.

Abbildung 2. Schematische Darstellung der Untersuchungshypothesen



Die ausformulierten Untersuchungshypothesen (H1 bis H10) lauten:

- H1: Die Struktur der transaktionalen Stresstheorie (s. Abbildung 2) lässt sich empirisch bestätigen.
- H2: Die prozessualen Annahmen der transaktionalen Stresstheorie spiegeln sich in Veränderungen in den Prozess- und Ergebnisvariablen und rekursiven Beziehungen zwischen den Modellkomponenten wieder. Veränderungen im psychischen Wohlbefinden und im Ausmaß an körperlichen Beschwerden wirken auf die Ereignis- und Ressourceneinschätzung zurück. Eine veränderte Ereignis- und Ressourceneinschätzung wiederum führt zu Veränderungen im Bewältigungsverhalten, was sich schließlich auf das psychische Wohlbefinden und das Ausmaß an körperlichen Beschwerden auswirkt.
- H3: Je höher die Neurotizismuswerte, desto negativer fällt die Ereigniseinschätzung aus (mehr Bedrohungs- und Verlusteinschätzungen, geringere Herausforderungseinschätzung) und desto geringer sind die situationsspezifischen Selbstwirksamkeitserwartungen ausgeprägt.

- H4: Je höher das Ausmaß an sozialer Unterstützung, desto positiver fällt die Ereigniseinschätzung aus (weniger Bedrohungs- und Verlusteinschätzungen, höhere Herausforderungseinschätzung) und desto höher sind die situationsspezifischen Selbstwirksamkeitserwartungen ausgeprägt.
- H5: In Abhängigkeit der Ereigniseinschätzung wird in einem unterschiedlichen Ausmaß auf problemzentrierte und emotionszentrierte Bewältigungsformen zurückgegriffen. Bei einer Verlust- und/oder Bedrohungsbewertung werden mehr problemzentrierte und emotionszentrierte Strategien genutzt, bei Herausforderungsbewertungen werden nur problemzentrierte Strategien verwendet. Höhere situationsspezifische Selbstwirksamkeitserwartungen gehen mit einer höheren Nutzung problemzentrierter und einer geringeren Nutzung emotionszentrierter Bewältigungsformen einher.
- H6: Die Haupteffekte der Ereignis- und Ressourceneinschätzung werden durch eine Interaktion zwischen beiden Bewertungsformen qualifiziert. Die Effekte der Bedrohungs- und Verlusteinschätzung auf das Bewältigungsverhalten werden durch situationsspezifische Selbstwirksamkeitserwartungen abgeschwächt, die die Herausforderungseinschätzung verstärkt.
- H7: Die Nutzung von Bewältigungsformen, sowohl problem- als auch emotionszentriert, hängt mit dem Ausmaß an körperlichen Beschwerden zusammen.
- H8: Die Nutzung von Bewältigungsformen, sowohl problem- als auch emotionszentriert, hängt mit dem psychischen Wohlbefinden zusammen.
- H9: Je höher die Neurotizismuswerte, desto geringer das wahrgenommene Ausmaß an sozialer Unterstützung.
- H10: Das Ausmaß an körperlichen Beschwerden und psychischem Wohlbefinden korrelieren negativ: Je geringer das Ausmaß an körperlichen Beschwerden, desto höher das Ausmaß an psychischem Wohlbefinden.

In zwei Studien an zwei unabhängigen Stichproben von LehramtsreferendarInnen wird die transaktionale Stresstheorie in der vorliegenden Arbeit überprüft. Studie 1 ist eine Querschnittuntersuchung an einer für Hessen repräsentativen Stichprobe von ReferendarInnen, Studie 2 ist eine Längsschnittuntersuchung mit insgesamt vier Messzeitpunkten, die sich auf einen Zeitraum von 13 Monaten verteilen. Ziel von Studie 1 ist zunächst die Überprüfung der Struktur im Querschnitt. Die Durchführung von Studie 2 kommt einer Forderung von Lazarus (2000, S. 668) nach, der die Wichtigkeit von Längsschnittuntersuchungen hervorhebt, da „...this kind of research allows researchers to identify psychological structures, that is, stable

personality dispositions (traits). It allows the identification of changes (processes) in psychological reactions over time and diverse conditions. Both are important.” Dies erfordert die Anwendung entsprechender Auswertungsmethoden, die allerdings bislang in Untersuchungen, deren Ziel die Überprüfung der transaktionalen Stresstheorie war, bis auf eine Ausnahme (Gunthert et al., 1999) nicht angewendet wurde (s. Kap.4.2). Dieses Manko soll mit Studie 2 aufgehoben werden, es wird auf eine Auswertungsmethode zurückgegriffen, die es ermöglicht, sowohl intraindividuelle Veränderungen über die Zeit und deren Determinanten als auch strukturelle Zusammenhänge zwischen verschiedenen Variablen und interindividuelle Varianz gleichzeitig zu berücksichtigen. Es wird auf ein relativ neues Verfahren zurückgegriffen, das so genannte Autoregressive-Latent-Trajectory (ALT)-Modell (Bollen & Curran, 2004; Curran & Bollen, 2001; für Anwendung in der allgemeinen Stressforschung z.B. Garst, Frese & Molenaar, 2000; Wills, Sandy, Yaeger, Cleary & Shinar, 2001), welches in Kapitel 7.2.1 ausführlich begründet und beschrieben wird. Mit Studie 1 wird auch eine Querschnittstudie realisiert, die eine Überprüfung von Prozessen nicht ermöglicht. Im Rahmen einer Querschnittbefragung kann eine größere Stichprobe befragt werden, die es einerseits erlaubt, generalisierte Aussagen zu treffen (Problem der Repräsentativität von Längsschnittstichproben durch zum Beispiel selektive Mortalität, s. Kap. 7.2.2) und andererseits es ermöglicht, bei einer ausreichend großen Stichprobe zwei Substichproben zu bilden, um ein empirisches Modell kreuzzuvalidieren (s. Kap. 6.2.3.1).

6 Studie 1

Zunächst wird für Studie 1 das methodische Vorgehen beschrieben (Kap. 6.1). Daran anschließend werden die Ergebnisse dargestellt (Kap. 6.2) und schließlich diskutiert (Kap. 6.3).

6.1 Methode Studie 1

6.1.1 Vorgehen

Ein standardisierter Fragebogen wurde an insgesamt 17 hessischen Studienseminaren verteilt, wovon acht Studienseminare ReferendarInnen für das Lehramt an Grund-, Haupt-, Real- und Sonderschulen ausbilden, sechs für das Lehramt an Gymnasien und drei für das Lehramt an beruflichen Schulen. Untersuchungszeitraum war Oktober und November 2001.⁴ Die Auswahl der Studienseminare erfolgte nach Maßgabe der Erreichbarkeit für den Untersuchungsdurchführenden. Es wurde darauf geachtet, dass die Studienseminare sowohl in ländlichen Gebieten als auch in Großstädten angesiedelt sind. Die Studienseminarleitung wurde in einem ersten Schritt schriftlich kontaktiert und über das geplante Forschungsprojekt informiert. Daran anschließend wurden Termine vereinbart und die Einwilligung des Seminar- und Personalrats der Studienseminare eingeholt. Insgesamt wurden 18 Studienseminare kontaktiert, von denen ein Studienseminar einer Befragungsdurchführung nicht einwilligte, da die Studienseminarleitung aufgrund einer bevorstehenden Zusammenlegung des Studienseminars mit einem anderen Seminar den ReferendarInnen keine weiteren Belastungen zumuten wollte.

Die Verteilung der Fragebogen erfolgte mehrheitlich über AusbilderInnen am Studienseminar, zum Teil konnten die Fragebogen aber auch im Rahmen von Vollversammlungen direkt an die Zielpersonen weitergegeben werden. Das zeitliche Fenster von zwei Monaten für die Befragungsdurchführung gewährleistet relativ ähnliche Bedingungen zwischen den Studienseminaren.

Zu Beginn der Befragung wurden alle ReferendarInnen darüber informiert, dass die Untersuchung in ein Forschungsprojekt eingebettet ist, welches sich mit Belastungen und Beanspruchungen im Referendariat beschäftigt. Es wurde die vollständig anonyme Auswertung der Daten zugesichert, die Teilnahme an der Befragung war freiwillig und wurde nicht

⁴ Bei der Erstkontaktierung der Studienseminare konnten nur 2 Studienseminare für das Lehramt an beruflichen Schulen erreicht werden. Aus diesem Grund wurde im März 2002 ein weiteres Studienseminar für das Lehramt an beruflichen Schulen kontaktiert und im April desselben Jahres konnten dort Daten nach erhoben werden. Neue Ausbildungszyklen beginnen in allen Studienseminaren im November und Mai eines jeden Jahres, also jedes halbe Jahr. Somit befanden sich die befragten ReferendarInnen dieses Studienseminars in derselben Phase ihrer Ausbildung wie die übrigen Befragten, nur eben sechs Monate später.

vergütet. Alle teilnehmenden ReferendarInnen wurden am Ende der Befragung über den genauen Forschungsanlass in Form eines schriftlichen Berichts aufgeklärt, indem auch die zentralen Ergebnisse dargestellt wurden.

6.1.2 Stichprobe

Die anvisierte Stichprobe sollte sich aus ReferendarInnen der Differenzierungsphase, Intensivphase I oder Intensivphase II zusammensetzen. Die Fragebogen von Befragten, die entweder keine Angabe zu der Ausbildungsphase machten oder bereits in der Prüfungsphase waren, wurden von den Analysen ausgeschlossen. Von den 659 zurückgegebenen Fragebogen mussten 15 Fragebogen aussortiert werden⁵, die Stichprobe umfasst also $N = 644$ ReferendarInnen. Potenziell erreichbar wären 1541 ReferendarInnen gewesen⁶, was einer Rücklaufquote von 42.54% entspricht, die als zufrieden stellend bezeichnet werden kann. Die Rücklaufquoten zwischen den Studienseminaren schwankt sehr stark (Min: 17%, Max: 83%; s. Tabelle A7 im Anhang, S. XXIV), was einerseits an der Methode der Fragebogenweitergabe (die direkte Verteilung der Fragebogen durch den Untersuchungsleiter erzielte eine höhere Rücklaufquote als die indirekte Verteilung der Fragebogen über AusbilderInnen), andererseits auch am unterschiedlichen Ausmaß der Kooperationsbereitschaft der Studienseminarleitungen lag.

Die Verteilung der Befragten auf die verschiedenen Lehrämter und das Geschlechtsverhältnis ist vergleichbar mit der Verteilung für Gesamthessen. In den Tabellen 2 und 3 sind die entsprechenden Angaben dargestellt.

⁵ Die 15 Befragten konnten nicht berücksichtigt werden, da sie entweder bereits die Examensprüfungen beendet hatten oder keiner Ausbildungsphase eindeutig zugeordnet werden konnten.

⁶ Für alle Studienseminare liegt die Anzahl der ReferendarInnen in den verschiedenen Ausbildungsphasen zum Zeitpunkt der Befragung vor.

Table 2. Verteilung der Befragten auf die verschiedenen Studienseminarformen im Vergleich zu Hessen

Lehramt	Anzahl	Prozentualer Anteil	Vergleichszahlen Hessen 2001¹
Grund-, Haupt-, Real- und Sonderschulen	316	49.1	53.2
Gymnasien	233	36.2	35.9
berufliche Schulen ²	9	14.8	11

Anmerkungen: 1 = Angaben des Hessischen Statistischen Landesamts 2001; 2 = einschließlich Fachlehreranwärter

Table 3. Geschlechtsverteilung der Befragten für die einzelnen Studienseminarformen im Vergleich zu Hessen

Lehramt	Befragte insgesamt	weiblich	Prozentualer Anteil	Vergleichszahlen Hessen 2001¹
Grund-, Haupt-, Real- und Sonderschulen	316	261	82.6	83.1
Gymnasien	233	136	58.4	59.1
berufliche Schulen ²	95	49	51.6	48.3

Anmerkungen: 1 = Angaben des Hessischen Statistischen Landesamts 2001; 2 = einschließlich Fachlehreranwärter

Die Verteilung der Befragten auf die verschiedenen Lehramter weicht statistisch signifikant von der Verteilung in Gesamthessen ab ($\chi^2 = 10.26$, $df = 2$, $p < .01$). Diese Abweichung resultiert vorwiegend aus einer Übergewichtung der Stichprobe hinsichtlich von ReferendarInnen aus Studienseminaren für das Lehramt an beruflichen Schulen. Diese Übergewichtung wurde vorgenommen, um bei der Auswertung der Daten Subgruppenanalysen durchführen zu können. Die Geschlechtsverteilung in der Stichprobe entspricht dagegen der für

Gesamthessen ($\chi^2 = 0.25$, $df = 1$, $p > .25$). Somit kann zumindest in Teilen von einer ausgewogenen Stichprobenziehung gesprochen werden.

Von den Befragten befanden sich zum Zeitpunkt der Befragungsdurchführung $N = 205$ in der Differenzierungsphase, $N = 218$ in der Intensivphase I und $N = 221$ in der Intensivphase II. Das Durchschnittsalter beträgt $M = 29.49$ Jahre ($SD = 4.19$, $Min = 22$, $Max = 45$), 69.9% leben in einer Partnerschaft, 20% haben ein oder mehrere Kinder.

6.1.3 Messinstrument

Als Messinstrument dient ein durch den Autor konzipierter Fragebogen, der alle relevanten Komponenten des Stressmodells erfasst (s. Kap. 3). In Tabelle 4 sind die Komponenten des Stressmodells bzw. Konstrukte und deren entsprechende Operationalisierung aufgeführt. Die Reihenfolge der Konstrukte in Tabelle 4 entspricht der im Fragebogen, wobei die Items zur Erfassung der Ereignis- und Ressourceneinschätzung untereinander in zufälliger Reihenfolge gemischt wurden.

Tabelle 4. Konstrukte und deren Operationalisierung

Konstrukt	Operationalisierung
Ausmaß der wahrgenommenen Belastungen (globales Stressausmaß)	Belastungsfragebogen (van Dick, 1999a), ergänzt um referendariatspezifische Belastungsfaktoren
Personale und situationale Ressourcen	Soziale Unterstützung (van Dick, 1999a), ergänzt um eigene Items Subskala „Neurotizismus“ des NEO-FFI (Borkenau & Ostendorf, 1993)
Ereigniseinschätzung („primary appraisal“)	Herausforderungs-, Bedrohungs- und Verlusteinschätzung von Lehrern (Schwarzer & Jerusalem, 1999)
Ressourceneinschätzung („secondary appraisal“)	Allgemeine Selbstwirksamkeitsskala (Schwarzer & Jerusalem, 1999) – Auswahl der 5 trennschärfsten Items Skala zur Lehrer-Selbstwirksamkeitserwartung (Schwarzer & Jerusalem, 1999) – Auswahl der 5 trennschärfsten Items
Bewältigungsstrategien („Coping“)	Coping (van Dick, 1999a), ergänzt um eigene Items
Gesundheit (körperliche Beschwerden)	Beschwerdefragebogen (van Dick, 1999a)
Psychisches Wohlbefinden	Fragebogen zum Gesundheitszustand (Bullinger & Kirchberger, 1998) – gekürzte Version

Darüber hinaus wurde noch eine stark gekürzte Version des AVEM (Arbeitsbezogenes Verhaltens- und Erlebensmuster; Schaarschmidt & Fischer, 1996), ein Fragebogen zur Messung organisationaler Identifikation (van Dick, Wagner, Stellmacher & Christ, 2004) und ein Messinstrument zur Erfassung von Organizational Citizenship Behavior in Anlehnung an van Dick (1999b) mit eingesetzt, die allerdings nicht Gegenstand der Fragestellung der vorliegen-

den Untersuchung sind. Aus diesem Grund wird auf diese Konstrukte nicht näher eingegangen.

Bei den ausgewählten Messinstrumenten handelt es sich um evaluierte Fragebogen, die sich bereits im Forschungseinsatz bewährt haben. Zum Teil wurden Modifikationen vorgenommen, um die Messinstrumente auf den Kontext der Befragung, das Lehramtsreferendariat, anzupassen. Auf die Modifikationen wird weiter unten näher eingegangen. Der Umfang des Fragebogens beträgt neun Seiten und die Bearbeitung hat im Durchschnitt etwa 20 bis 30 Minuten benötigt. Der komplette Fragebogen ist im Anhang aufgeführt (s. S. IIIff).

6.1.3.1 Aufbau des Fragebogens

Im Einleitungsteil wurden einige soziodemographische und mit der Ausbildung zusammenhängende Angaben abgefragt wie z.B. das Geschlecht, Alter, Familienstand, Kinderlosigkeit (Kinder ja/nein), angestrebtes Lehramt, Größe der Ausbildungsschule etc. Des Weiteren wurde mit jeweils einer Frage erhoben, wie zufrieden die ReferendarInnen mit der Ausbildung an ihrem Studienseminar und der Anleitung in der jeweiligen Ausbildungsschule sind. Zusätzlich enthielt der Fragebogen noch Fragen zur Identifikation der ReferendarInnen mit der eigenen Karriere, der Ausbildungsschule und der Berufsgruppe der Lehrer, zum Extra-Rollen-Verhalten (Organizational Citizenship Behavior) und zum arbeitsbezogenen Erlebens- und Verhaltensmuster (s.o.).

Im Folgenden wird auf die eingesetzten Skalen näher eingegangen.

Belastungsmaß. Hier sollten die Befragten einschätzen, wie sehr sie sich aktuell durch eine Reihe von Belastungsfaktoren, die im Schulalltag bzw. während des Referendariats auftreten, belastet fühlen. Insgesamt wurden 23 mögliche Belastungsfaktoren aufgelistet. Die Belastungsfaktoren lassen sich in zwei Kategorien unterteilen (s. Kap. 5.1): Zum einen in Schul- und berufsbezogene Belastungsfaktoren und zum anderen in ausbildungsbezogene Belastungsfaktoren.

Schul- und berufsbezogene Belastungsfaktoren wurden aus einer Studie von van Dick (1999a) entnommen, insgesamt werden 13 Faktoren aufgelistet (s. Tabelle 5). Van Dick orientierte sich bei der Auswahl der Faktoren an Studien zu Belastungen im Lehrerberuf. Die Belastungsfaktoren ließen sich an einer großen Lehrerstichprobe zu einem Faktor zusammenfassen, die interne Konsistenz für die Gesamtskala war mit $\alpha = .79$ zufrieden stellend (van Dick, 1999a, S. 128).

Zur Auswahl ausbildungsbezogener Belastungsfaktoren wurden Interviews mit einem Studienseminarleiter und mehreren ehemaligen Referendarinnen durchgeführt. Auf Basis der

Angaben der Interviewten wurden insgesamt 10 Belastungsfaktoren ausgewählt, die in Tabelle 5 aufgelistet sind.

Tabelle 5. Lehrertypische und ausbildungsspezifische Belastungsfaktoren (Reihenfolge wie im Fragebogen)

Quelle	Belastungsfaktoren
<i>Schul- und berufsbezogen</i>	Zu große Klassen (belast01)
	Mangelnde Motivation bzw. Konzentrationsfähigkeit der SchülerInnen (belast02)
	Disziplinprobleme (belast03)
	Probleme mit ausländischen SchülerInnen (belast04)
	Ärger mit Behörden bzw. Institutionen (belast05)
	Probleme mit den Eltern (belast06)
	Fachfremder Unterrichtseinsatz (belast07)
	Unterschiedliche Lernvoraussetzungen der SchülerInnen (belast08)
	Vor- und Nachbereitung des Unterrichts (belast09)
	Geringe Lernbereitschaft von SchülerInnen (belast10)
	Verwaltungsarbeit (belast11)
	Ständige Kritik am Lehrerberuf und fehlende Anerkennung in der Öffentlichkeit (belast12)
	Hektisches und störungsvolles Arbeitsklima an der Schule (belast13)
<i>Ausbildungsbezogen</i>	Unterrichtsbesuche / Lehrproben (belast14)
	Konflikte mit der Ausbildungsschule (belast15)
	Eigenverantwortlicher Unterricht (belast16)
	Konflikte mit AusbilderInnen (belast17)
	Pädagogische Prüfungsarbeit (belast18)
	Examensprüfungen (belast19)
	Klassenfahrten (belast20)
	Seminarveranstaltungen am Studienseminar (belast21)
	Einstellungsdruck (belast22)
	Unterrichtsentwürfe (belast23)

Die Belastungsfaktoren konnten auf einer sechsstufigen Ratingskala mit den Endpolen 1 = „überhaupt nicht belastend“ und 6 = „sehr belastend“ beantwortet werden, höhere Werte bedeuten ein höheres Ausmaß subjektiver Belastung.

Neurotizismus. Neurotizismus wurde mit der Neurotizismus-Subskala der deutschen Übersetzung des NEO-FFI (Borkenau & Ostendorf, 1993; s. a. Borkenau & Ostendorf, 1991; Original: Costa & McCrae, 1989) erfasst. Die Subskala umfasst 12 Items (z.B. „Ich fühle mich oft angespannt und nervös“; Gesamtskala s. Anhang, S. XIX), die interne Konsistenz in

einer kürzlich erfolgten Validierung des NEO-FFI an einer deutschen Bevölkerungsstichprobe (Körner, Geyer & Brähler, 2002) schwankt zwischen $\alpha = .81 - .86$ und kann als zufrieden stellend bezeichnet werden. Die faktorielle Struktur des NEO-FFI wird in der Studie von Körner et al. (2002) als problematisch angesehen, allerdings weisen die Autoren daraufhin, dass die Neurotizismus-Subskala relativ unproblematisch sei. Da das Originalitem „Ich bin nicht leicht beunruhigt“ oft zu Irritationen bei Probanden führt, wurde es in der vorliegenden Untersuchung leicht abgewandelt in „Ich bin leicht beunruhigt“. Laut Körner et al. (2002) sprechen die Korrelationen der NEO-FFI-Skalen mit einem Messinstrument zur Erfassung von Selbstkonzeptmerkmalen (Gießen-Test; Beckmann, Brähler & Richter, 1991) für die Konstruktvalidität des Verfahrens. So korreliert die Neurotizismus-Subskala des NEO-FFI erwartungsgemäß negativ mit der Subskala soziale Resonanz und positiv mit der Subskala depressive Grundstimmung des Gießen-Test.

Das Problem der Verfälschungssensitivität von Persönlichkeitstests und der damit verbundenen Validitätsminderung, was auch für den NEO-FFI gilt, ist bekannt (s. Krahe & Herrmann, 2003), es kann aber davon ausgegangen werden, dass im vorliegenden Untersuchungskontext Antwortverzerrungen eher eine untergeordnete Rolle spielen, da es sich nicht um eine Auswahl-situation wie in der Personalauswahl handelt und das Ausfüllen der Items freiwillig war. Im Original sollen die Items des NEO-FFI auf einer fünfstufigen Ratingskala von 0 = „starke Ablehnung“ bis 4 „starke Zustimmung“ beantwortet werden. Für die vorliegende Untersuchung wurde das Antwortformat an das der anderen Skalen angepasst⁷, so dass ebenfalls eine sechsstufige Ratingskala zur Beantwortung zur Verfügung stand von 1 = „trifft nicht zu“ bis 6 = „trifft sehr zu“. Höhere Werte bedeuten eine höhere Neurotizismusausprägung.

Ausmaß sozialer Unterstützung. Das Ausmaß an sozialer Unterstützung wurde in Anlehnung an van Dick (1999a) in Form einer Tabelle erfasst (s. Tabelle 6). Dabei sollten die Versuchspersonen für sechs Bereiche (AusbilderInnen, KollegInnen, Privatbereich und SchülerInnen) das Ausmaß an sozialer Unterstützung einschätzen, zum einen hinsichtlich praktischer, zum anderen hinsichtlich emotionaler Unterstützung, jeweils bezogen auf Probleme im Zusammenhang mit dem Referendariat. Hierzu sollten die Befragten Zahlen in die jeweils dafür vorgesehene Zelle der Tabelle zwischen 1 = „überhaupt keine Unterstützung“ bis 6 = „völlige Unterstützung“ eintragen. Höhere Werte bedeuten ein höheres Ausmaß an sozialer Unterstützung.

⁷ Da keine Normvergleiche in der vorliegenden Untersuchung vorgesehen sind, ist die Änderung des Antwortformats unproblematisch.

Tabelle 6. Messinstrument zur Erfassung sozialer Unterstützung

1 = überhaupt keine Unterstützung 6 = völlige Unterstützung	Ausbilder- Innen	Mentoren- Innen	Referendars- kollegInnen	KollegInnen an der Ausbildungs- schule	Partner oder Partnerin und Freunde	Schüler
Wie sehr werden Sie von folgenden Personen praktisch unterstützt, wenn die Arbeit im Zusammenhang mit dem Referendariat schwierig wird?						
Wie sehr fühlen Sie sich von folgenden Personen emotional unterstützt, wenn Probleme im Zusammenhang mit dem Referendariat auftauchen?						

Die gewählte tabellarische Form erlaubt eine ökonomische Messung des Ausmaßes an sozialer Unterstützung und hat sich in verschiedenen Studien bewährt (Schnitger & van Dick, 2000; van Dick, 1999a).

Ereigniseinschätzung. Die Ereigniseinschätzung im Sinne der primären Bewertung wurde mit einer von Jerusalem (1999) entwickelten Skala erfasst. Die Skala setzt sich aus den Subskalen Herausforderung (2 Items, z.B. „Ich bin zuversichtlich, dass ich den Anforderungen im Beruf gewachsen bin“), Bedrohung (3 Items, z.B. „Ich befürchte, dass mich die beruflichen Belastungen überfordern“) und Verlust (3 Items, z.B. „Ich bin niedergeschlagen, weil meine berufliche Situation so schlecht ist“) zusammen, die Skalen wurden in verschiedenen Studien empirisch geprüft (s. Jerusalem, 1991, 1993; Schwarzer, 2000) und weisen in verschiedenen Stichproben eine zufrieden stellende interne Konsistenz auf. Das Antwortformat in der Originalfassung ist vierstufig von 1 = „stimmt kaum“ bis 4 = „stimmt genau“, auch hier wurde es an die übrigen eingesetzten Skalen angepasst, so dass den Befragten zur Beantwortung ebenfalls eine sechsstufige Ratingskala von 1 = „trifft nicht zu“ bis 6 = „trifft sehr zu“ zur Verfügung stand. Höhere Werte bedeuten eine höhere Ausprägung in den entsprechenden Skalen. Die Skala kann dem Anhang entnommen werden (S. XXI).

Selbstwirksamkeitserwartungen. Zur Operationalisierung der Selbstwirksamkeitserwartungen wurde zum einen die revidierte Fassung der Allgemeinen Selbstwirksamkeitsskala von Schwarzer und Jerusalem (1999) und zum anderen die Lehrer-Selbstwirksamkeitsskala von Schmitz und Schwarzer (2000), die situationsspezifische Selbstwirksamkeitserwartungen erfasst, eingesetzt. Bei beiden Skalen handelt es sich um hoch reliable und valide Skalen

(Schmitz, 1999; Schmitz und Schwarzer, 2000; Scholz, Gutierrez-Dona, Sud & Schwarzer, 2002). Jede Skala umfasst in den Originalversionen 10 Items, aus ökonomischen Erwägungen wurden jeweils nur die fünf trennschärfsten Items (Beispielitem Allgemeine Selbstwirksamkeitserwartungen: „Wenn sich Widerstände auftun, finde ich Mittel und Wege, mich durchzusetzen“; Beispielitem Lehrer-Selbstwirksamkeitserwartungen: „Ich bin mir sicher, dass ich auch mit den problematischen Schülern in guten Kontakt kommen kann, wenn ich mich darum bemühe“) ausgewählt, basierend auf den Angaben in der Skalendokumentation im Rahmen der wissenschaftlichen Begleitung des Modellversuchs „selbstwirksame Schulen“ (Schwarzer & Jerusalem, 1999). Die Items konnten auf einer sechsstufigen Ratingskala von 1 = „trifft nicht zu“ bis 6 = „trifft genau zu“ beantwortet werden, höhere Werte bedeuten eine höhere Ausprägung in dem entsprechenden Merkmal. Die komplette Skala mit Zuordnung zu der entsprechenden Originalskala kann dem Anhang entnommen werden (S. XXI).

Bewältigungsstrategien (Coping). Zur Erfassung unterschiedlicher Bewältigungsstrategien wurde eine Copingskala von van Dick (1999a) eingesetzt, die speziell für den Lehrerberuf entwickelt wurde. Bei der Generierung der Items der Skala zog der Autor zum Teil den Stressverarbeitungsfragebogen von Janke, Erdman und Kallus (1995) und Studien zu Lehrerstress heran (z.B. Kyriacou & Pratt, 1985; Dewe, 1985), zum anderen Teil wurden auch neue Items formuliert. Insgesamt wurden 33 Items einer empirischen Prüfung unterzogen (s. van Dick, 1999a, Studie 2), Faktoren- und Reliabilitätsanalysen führten zur Auswahl von 14 Items, die die folgenden Subdimensionen bilden: Situationskontrollversuche („Ich versuche, die Situation aktiv zu verändern.“), Suche nach sozialer Unterstützung („Ich spreche mit Freunden und Bekannten über die Schwierigkeiten.“), Resignation („Ich neige dann schnell dazu zu resignieren.“), Ignorieren („Ich fahre so schnell wie möglich nach Hause, um mich dort von den Problemen in der Schule zu erholen.“), Aggression („Ich ärgere mich über die Schulleitung und/oder die Kollegen.“). Zur Auswertung verwendete van Dick (1999a) nur die Gesamtskala, da die Unterskalen über zu geringe interne Konsistenzen verfügten. Dieser Gesamtscore kann als bipolare Coping-Dimension mit den Endpolen emotionszentriertem und problemzentriertem Bewältigungsverhalten interpretiert werden. Für die vorliegende Untersuchung wurden zusätzlich zu den 14 Items der Copingskala noch drei weitere Items formuliert, die sich auf den referendariatspezifischen Kontext beziehen: (a) „Ich suche mir andere ReferendarInnen, die mir bei der Lösung der Probleme helfen.“, (b) „Ich bitte die AusbilderInnen an meinem Studienseminar um Hilfe.“, (c) „Ich hole mir Rat bei meinem/meiner MentorIn.“. Die Probanden wurden instruiert, bei der Beantwortung der Items an Situationen zu denken, die sie im Referendariat aufregen und belasten. Alle Items

konnten mittels einer sechsstufigen Ratingskala von 1 = „trifft nicht zu“ bis 6 = „trifft genau zu“ beantwortet werden. Die komplette Skala kann im Anhang eingesehen werden (S. XXII).

Körperliche Beschwerden. Körperliche Beschwerden wurden mit einer gekürzten Version des Gießener Beschwerdefragebogens (Brähler & Scheer, 1983) erfasst, welcher in der hier eingesetzten Version von van Dick (1999a; s.a. Lenke, 1997) konzipiert und empirisch geprüft wurde. Auf Basis teststatistischer Kennwerte wurde die Originalanzahl von 57 Items auf 8 Items reduziert, wobei die ursprünglichen vier Dimensionen mit je zwei Items (Beispielitems: Dimension Magenbeschwerden „Sodbrennen oder saures Aufstoßen“; Dimension Erschöpfungsneigung „Mattigkeit“; Dimension Gliederschmerzen „Nacken und Schulterschmerzen“; Dimension Herzbeschwerden „Herzklopfen, Herzschnagen oder Herzschnalzen“) repräsentiert sind (für Details s. Lenke, 1997). Zusätzlich wurden noch „Stimmprobleme“ als ein weiteres lehrertypisches Symptom aufgenommen (Simberg, Laine, Sala & Rönnekaa, 2000; Smith, Lemke, Taylor, Kirchner & Hoffman, 1998). Auf einer sechsstufigen Skala von 1 = „nie“ bis 6 = „sehr häufig“ sollten die ReferendarInnen angeben, wie häufig sie unter den aufgelisteten Beschwerden im vergangenen Monat litten. Höhere Werte zeigen ein höheres Ausmaß an körperlichen Beschwerden an. Die komplette Skala kann im Anhang eingesehen werden (S. XXIII).

Psychisches Wohlbefinden. Zur Messung psychischen Wohlbefindens wurde auf die gleichnamige Subskala der deutschen Version (Bullinger & Kirchberger, 1998) des Fragebogens zum Gesundheitszustand (SF-36; McHorney, Ware & Raczek, 1993; McHorney, Ware, Lu & Sherbourne, 1994; Ware & Sherbourne, 1992) zurückgegriffen. Der SF-36 hat als Ziel, den subjektiven Gesundheitszustand bzw. die gesundheitsbezogene Lebensqualität zu erfassen. Das Verfahren kann hinsichtlich teststatistischer Gütekriterien (Bullinger; Kirchberger & Ware, 1995) als vorbildlich bezeichnet werden, auch eine sorgfältige Validierung (Bullinger, 1995) wurde durchgeführt und weist auf die Gültigkeit des Verfahrens hin. Insgesamt werden in der Subskala „Psychisches Wohlbefinden“ acht Indikatoren aufgelistet, die Befragten sollen auf einer sechsstufigen Ratingskala von 1 = „immer“ bis 6 = „nie“ angeben, wie oft sie sich im vergangenen Monat entsprechend gefühlt haben (z.B. „voller Schwung“; „ruhig und gelassen“). Alle Contrasts wurden für die Auswertung umgepolt, so dass höhere Werte ein höheres Ausmaß an psychischem Wohlbefinden anzeigen. Die komplette Skala kann im Anhang eingesehen werden (S. XXIII).

6.2 Ergebnisse Studie 1

Zunächst wird auf das allgemeine Vorgehen bei der Datenanalyse eingegangen (Kap. 6.1.1). Anschließend werden das Datenscreening (Kap. 6.1.2) und die Datenaufbereitung (Kap. 6.1.3) beschrieben. Im nächsten Abschnitt werden die eingesetzten Skalen überprüft (Kap. 6.1.4) und schließlich das Messmodell analysiert (Kap. 6.1.5). Nach Überprüfung des Stressmodells (Kap. 6.1.6) wird der Ergebnisteil von einem zusammenfassenden Fazit der Ergebnisse (Kap. 6.1.8) abgeschlossen.

6.2.1 Allgemeines Vorgehen

Ziel der vorliegenden Studie ist es, das im Theorieteil vorgestellte theoretische Stressmodell empirisch in seiner Gültigkeit zu überprüfen (s. Kap. 5). Ein Verfahren zur Modellierung kausaler Strukturen bei querschnittlich erhobenen Daten stellen Strukturgleichungsanalysen (SEM) dar (für einen Überblick s. MacCallum & Austin, 2000; Marcoulides & Schumacker, 2001; Ulman & Bentler, 2003).⁸ Hauptvorteil dieser Methode ist, dass der Messfehler bei der Schätzung von Modellparametern berücksichtigt wird und komplexe Beziehungen zwischen Variablen geschätzt werden. An dieser Stelle soll aber bereits darauf hingewiesen werden, dass mit Hilfe von SEM keine kausalen Beziehungen zwischen Modellvariablen überprüft werden können (Nachtigall, Kroehne, Funke & Steyer, 2003). Hierzu sind experimentelle Studien und/oder Längsschnittuntersuchungen notwendig. Vom Prinzip her wird in SEM überprüft, ob ein theoretisch postuliertes Modell (mit seinen kausalen Beziehungen) eine geschätzte Populations-Kovarianzmatrix reproduzieren kann, die konsistent ist mit einer (beobachteten) Kovarianzmatrix einer gegebenen Stichprobe. Als Leitlinien für die Bewertung der Übereinstimmung zwischen der geschätzten und der empirischen Kovarianzmatrix stehen verschiedene Fit-Indizes zur Verfügung. Die klassische Form, den Fit eines Modells abzuschätzen, stellt die χ^2 -goodness-of-fit-Statistik dar. Dabei wird die Diskrepanz zwischen der Stichproben- und der gefitteten Kovarianzmatrix gemessen. Problem dieser Statistik ist u.a., dass sie stark abhängig von der Stichprobengröße ist: Je größer die Stichprobe, desto wahrscheinlicher werden passende Modelle abgelehnt (Hu & Bentler, 1995). Aus diesem Grund wurden eine Reihe von Fit-Indizes vorgeschlagen, die die Nachteile der χ^2 -goodness-of-fit-Statistik beheben sollen. Allerdings werden auch diese Fit-Indizes durch die Stichprobengröße, das Schätzverfahren und die Verletzung der Verteilungsannahmen und Varianzabhängigkeiten beeinflusst (Hu & Bentler, 1995). Somit müssen auch diese Fit-Indizes vorsichtig

⁸ Zur Berechnung von SEM wird das Programm AMOS 5.0 (Arbuckle, 2003; Arbuckle & Wothke, 1999) verwendet.

interpretiert werden, wenn auch bestimmte Cutt-off Werte für die Indizes vorgeschlagen werden, die anzeigen sollen, ab wann ein Wert auf ein gut angepasstes Modell hinweist.

In der vorliegenden Arbeit wird auf die von Hu und Bentler (1999; s.a. Schermelleh-Engel, Moosbrugger & Müller, 2003) vorgeschlagenen Indizes und Cutt-off Werte zurückgegriffen. Die Indizes sind der Comparative Fit Index (CFI; Bentler, 1990)⁹, der root mean squared error of approximation (RMSEA; Steiger & Lind, 1980)¹⁰ und der standardized root mean squared residual (SRMR; Bentler, 1995).¹¹ Hu und Bentler (1999; vgl. Marsh, Hau & Wen, 2004) kommen, basierend auf einer Monte Carlo Studie, zu der Empfehlung, dass zur Bewertung eines Modells mindestens zwei Indizes herangezogen werden sollten, da in Hinblick auf unterschiedliche Probleme (z.B. Missspezifikation des Modells, kleine Stichprobengröße) verschiedene Indizes unterschiedlich effizient sind und gängige Cutt-off Werte falsch sein können. Bei Nutzung mehrerer Indizes kann solchen Problemen vorgebeugt werden. Als Cutt-off Werte schlagen Hu und Bentler für den CFI einen Wert nahe .95, für den RMSEA einen Wert von ungefähr .06 und für den SRMR einen Wert um die .08 vor. Zur Vollständigkeit wird auch die χ^2 -goodness-of-fit-Statistik sowie das Verhältnis von χ^2/df angegeben. Dieses Verhältnis sollte idealerweise 2:1 sein, akzeptabel ist auch ein 3:1 Verhältnis (Marsh & Hocevar, 1985).

Browne und Cudeck (1993) weisen darüber hinaus auf das Problem von Punktschätzungen von Fit-Indizes hin. Auch hier kann nicht, wie bei Parameterschätzungen im Allgemeinen, davon ausgegangen werden, dass diese Schätzung für die Population zutrifft. So können Modelle mit akzeptablen Fit-Werten irrtümlich angenommen werden, da die Schätzer der Fit-Indizes fehlerhaft sind. Browne und Cudeck empfehlen daher, Konfidenzintervalle für die Parameterschätzer anzugeben. Für den RMSEA ist dies in AMOS 5.0 implementiert. Weiterhin empfehlen Browne und Cudeck (vgl. MacCallum, Browne & Sugawara, 1996) den so genannten Test für close-fit zu verwenden. Dabei wird die Nullhypothese überprüft, ob der RMSEA in der Population kleiner als .05 ist. Als akzeptabel werden dabei Werte von $.05 < p \leq .10$, als gut werden Werte von $.10 < p \leq 1.00$ angesehen (Schermelleh-Engel et al.,

⁹ Der CFI gehört zu den inkrementellen Fit-Indizes, die die proportionale Steigerung des Fits im Vergleich zu einem restriktiveren, genesteten Vergleichsmodell messen. Meist dient als Vergleichsmodell ein so genanntes Nullmodell, in dem alle beobachteten Variablen unkorreliert sind. Der CFI schwankt zwischen 0 und 1, höhere Werte geben einen besseren Fit an.

¹⁰ Der RMSEA gehört zu der Gruppe der absoluten Fit-Indizes, denen gemeinsam ist, dass sie anzeigen, wie gut ein Modell die Stichprobendaten reproduziert. Dabei wird kein Vergleichsmodell herangezogen, um die Verbesserung bzw. die Steigerung im Modell-Fit zu schätzen. Allerdings wird meist ein impliziter bzw. expliziter Vergleich mit einem gesättigten Modell vorgenommen. Der RMSEA kann zwischen 0 und 1 schwanken, kleinere Werte entsprechen einem besser Fit.

¹¹ Der SRMR gehört ebenfalls zu den absoluten Fit-Indizes. Er schwankt ebenfalls zwischen 0 und 1, wie beim RMSEA indizieren kleinere Werte einen besseren Fit.

2003). In der vorliegenden Arbeit wird für alle Endmodelle (Mess- und Strukturmodell) zusätzlich zu den oben genannten Fit-Indizes der p-close-Fit angegeben.

Als statistisches Verfahren zur Modellschätzung und Modellbewertung wird das Maximum-Likelihood-Schätzverfahren (ML) verwendet (Browne, 1982; Jöreskog, 1967), als Input dienen Kovarianzmatrizen.

6.2.2 Datenscreening

Das Screening der Daten für Studie 1 erfolgte in mehreren Schritten. Zunächst wurden die Daten sorgfältig auf Fehler bei der Eingabe geprüft und gegebenenfalls korrigiert. In einem zweiten Schritt wurde für jede Versuchsperson die Anzahl der fehlenden Werte in den zentralen Variablen (d.h. alle Variablen, die zu den Skalen gehören, die die Konstrukte des Stressmodells erfassen, sowie den Belastungsitems; s. Kap. 6.1.3) ausgezählt. Befragte mit mehr als 10% fehlender Werte wurden von den weiteren Analysen ausgeschlossen (s. Kline, 1998), da bei diesen Versuchspersonen meist ganze Skalen nicht ausgefüllt wurden. Insgesamt erfüllten acht Befragte dieses Kriterium und wurden somit von den weiteren Analysen ausgeschlossen. Die endgültige Analytestichprobe umfasst demnach $N = 636$ ReferendarInnen.

Bei diesem Analyseschritt zeigt sich darüber hinaus, dass bei zwei Items der Sozialen Unterstützungsskala und einem Item der Copingskala überzufällig viele fehlende Werte zu verzeichnen sind, die nur bei Befragten aus drei der teilnehmenden Studienseminare auftreten¹². Alle drei Items haben Aussagen zu den MentorInnen an den Ausbildungsschulen zum Thema. Nach Rücksprache mit den entsprechenden Studienseminaren ist dies auf ein fehlendes MentorInnen-System zurückzuführen. Um die Vergleichbarkeit zu gewährleisten, wurden diese drei Items bei allen weiteren Analysen nicht mehr berücksichtigt.

6.2.3 Datenanalyse - Vorschritte

Vor den eigentlichen Analysen wurden die Daten einigen Vorschritten unterzogen. Zunächst wurde die Gesamtstichprobe in zwei Substichproben unterteilt. Anschließend wurden fehlende Werte jeweils in beiden Substichproben ersetzt und die Daten auf Normalverteilung und multivariate Ausreißer geprüft.

¹² Der prozentuale Anteil fehlender Werte in den drei Variablen liegt für die angesprochenen Studienseminare bei $M = 26.83\%$, für alle anderen Studienseminare bei $M = 0.99\%$.

6.2.3.1 Aufteilung der Gesamtstichprobe in 2 Substichproben

Zur Überprüfung des Stressmodells wird dabei zunächst an einer Hälfte der Stichprobe (Kalibrierungsstichprobe) das Modell mittels Strukturgleichungsanalysen überprüft und gegebenenfalls modifiziert. An der anderen Hälfte der Stichprobe (Validierungsstichprobe) wird dann anschließend das empirisch gewonnene Modell validiert. Dieses Vorgehen schützt in Teilen vor Stichprobenartefakten, d.h. dass die Ergebnisse auf Spezifika der Stichprobe zurückzuführen sind, die nicht kontrolliert werden konnten (Cudeck & Browne, 1983; MacCallum & Austin, 2000). Zu diesem Zweck wurde die Gesamtstichprobe in zwei Zufallsstichproben geteilt, wobei sowohl die Geschlechtsverteilung als auch die Anteile der unterschiedlichen Ausbildungsphasen (Differenzierungsphase, Intensivphasen I und II) und Studienseminare (Studienseminare für Berufsschulen, für Grund-, Haupt-, Real- und Sonderschulen und für Gymnasien) kontrolliert wurden¹³, um möglichen Verzerrungen bei der Zufallsaufteilung vorzubeugen (s. Tabelle 7).

Tabelle 7. Geschlechtsverteilung und Aufteilung auf die unterschiedlichen Ausbildungsphasen und Studienseminare (in Klammern in Prozent) in der Kalibrierungs- und Validierungsstichprobe

		Kalibrierungsstichprobe	Validierungsstichprobe
		(N = 318)	(N = 318)
Geschlecht	Männlich	97 (30.5)	98 (30.8)
	Weiblich	221 (69.5)	220 (69.2)
Ausbildungsphase	Differenzierungsphase	109 (34.3)	108 (34)
	Intensivphase I	108 (34)	108 (34)
	Intensivphase II	101 (31.8)	102 (32.1)
Studienseminar	Berufliche Schulen	48 (15.1)	47 (14.8)
	GHR	114 (35.8)	114 (35.8)
	Gymnasien	156 (49.1)	157 (49.4)

¹³ Die Anteile der Ausbildungsphasen, Studienseminare und des Geschlechts wurden für beide Zufallsstichproben konstant gehalten, die Wahrscheinlichkeit für jeden Befragten, der einen oder der anderen Zufallsstichprobe zugeordnet zu werden, blieb hiervon unberührt.

Auch in anderen soziodemographischen Variablen zeigen sich keinerlei statistisch bedeutsame Unterschiede zwischen der Kalibrierungs- und Validierungsstichprobe (s. Tabelle 8).¹⁴

Tabelle 8. Vergleich der beiden Teilstichproben in Hinblick auf weitere soziodemographische Variablen

		Kalibrierungsstichprobe		Validierungsstichprobe	
		(N = 318)		(N = 318)	
Familienstand	Ledig	100 (31.4)		88 (27.7)	
	Partner/Partnerin	218 (68.6)		230 (72.3)	
Kinder	Nein	258 (81.1)		247 (77.7)	
	Ja	59 (18.6)		70 (22)	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Alter		29.34	4.24	29.68	4.18
Größe der Ausbildungsschule (Anzahl der SchülerInnen)		805.31	646.80	850.48	663.66
Größe des Kollegiums an der Ausbildungsschule		54.63	36.26	57.17	36.76
Durchschnittliche Klassengröße		21.87	6.14	22.36	5.53

Die Ergebnisse des Vergleichs der beiden Zufallsstichproben hinsichtlich wichtiger soziodemographischer Variablen zeigen, dass beide Stichproben vergleichbar sind.

6.2.3.2 Ersetzen der fehlenden Werte

Im nächsten Schritt wurden in beiden Teilstichproben alle fehlenden Werte ersetzt. Hierzu stehen verschiedene Methoden zur Verfügung, üblich sind das paarweise (pairwise deletion, PD) oder listenweise (listwise deletion, LD) Löschen von Fällen mit fehlenden Werten in relevanten Variablen oder das Ersetzen der fehlenden Werte durch den Mittelwert aller beobachteten Werte (mean imputation, MI). Verschiedene Studien (Brown, C.H., 1983; Brown, R.L., 1994; Little & Rubin, 1987) zeigen allerdings, dass diese Verfahren zu verzerrten und/oder ineffizienten Schätzern und zu einem erhöhten Risiko indefiniter Stichproben-

¹⁴ Zur Überprüfung möglicher Unterschiede wurde für kategoriale abhängige Variablen χ^2 -Tests und für alle quantitativen Variablen eine einfaktorielle ANOVA durchgeführt.

Kovarianz-Matrizen führen. Darüber hinaus spielt zusätzlich eine Rolle, ob die fehlenden Werte zufällig zustande kommen oder ob den fehlenden Werten eine nicht zufällige Systematik zugrunde liegt. Nach Rubin (1976; Little & Rubin, 1987) werden Daten als missing completely at random (MCAR) bezeichnet, wenn die Antwortrate unabhängig von der Ausprägung des Merkmals selbst als auch von der Ausprägung anderer Merkmale ist. Eine etwas gelockerte Bedingung unterliegt Daten, die missing at random (MAR) aufweisen. Hier muss die Antwortrate nur unabhängig von der Ausprägung des Merkmals selbst sein. Schließlich gibt es noch observed at random (OAR), wobei dann die Antwortrate unabhängig von der Ausprägung anderer Merkmale ist. Little und Schenker (1995) kommen zu dem Ergebnis, dass PD, LD und MI auch unter MCAR zu unbefriedigenden Ergebnissen führen. PD- und LD-Schätzer sind zwar konsistent, aber ineffizient. MI führt zu verzerrten Varianz- und Kovarianz-Schätzern. Wenn die Daten MAR sind, liefern auch PD und LD verzerrte Schätzer. Auch Wothke (2000) kommt in einer Simulationsstudie zu gleichen Resultaten.

Alternativ zu bislang vorgestellten Methoden zum Umgang mit fehlenden Werten gibt es maximum likelihood (ML) Algorithmen (zur Übersicht s. Enders, 2001), die sowohl bei MCAR als auch MAR zu robusten Schätzern kommen und in verschiedenen Untersuchungen ihre Überlegenheit gegenüber den konventionellen Methoden zeigen (Arbuckle, 1996; Enders & Bandalos, 2001; Graham, Hofer & MacKinnon, 1996; Schafer & Graham, 2002; Wothke, 2000). Momentan stehen in üblichen Software-Paketen drei Algorithmen zur Verfügung: (1) der Multigruppen-Ansatz, (2) „full information maximum likelihood estimation“ (FIML) und (3) der „expectation maximization (EM) Algorithmus“ (zur Übersicht s. Enders, 2001). Die drei Methoden kann man weiter in direkte und indirekte Methoden unterteilen: Sowohl der Multigruppen-Ansatz als auch FIML sind direkte Ansätze, da lineare Modellparameter direkt aus den zur Verfügung stehenden Rohdaten ohne vorherige Präparation der Daten ermittelt werden. Im Multigruppen-Ansatz wird die Stichprobe in G Substichproben unterteilt in der Art und Weise, dass jede Substichprobe das gleiche Muster fehlender Werte aufweist. Eine likelihood-Funktion für jede der G Gruppen wird berechnet, diese gruppenweise likelihood-Funktionen werden über alle Gruppen akkumuliert und maximiert (d.h. die fallweise likelihood der beobachteten Daten wird maximiert). Der FIML-Ansatz geht ähnlich vor, allerdings wird hier die likelihood-Funktion auf der individuellen Ebene, nicht gruppalen Ebene berechnet. Somit besteht der Unterschied nur darin, dass die zu bestimmende Funktion die Summe von n paarweisen likelihood-Werten ist, während die Funktion des Multigruppen-Ansatzes die Summe von G gruppenweisen likelihood-Werten ist. Dagegen handelt es sich bei dem EM Algorithmus um einen indirekten Ansatz, hier müssen vor der Berechnung von

linearen Parameter-Schätzern zusätzliche Analysen vorgeschaltet werden. Der EM-Algorithmus liefert einen ML-Schätzer der Kovarianz-Matrix und des Mittelwertvektors, der wiederum als Input für weitere Modellierungen verwendet werden kann (wie z.B. SEM).

Enders (2001) diskutiert ausführlich Vor- und Nachteile der hier beschriebenen ML-Methoden. Ein Hauptproblem des EM Algorithmus sind mögliche verzerrte Standardfehler, was gerade bei Strukturgleichungsanalysen problematisch sein kann (s.a. Graham & Schafer, 1999), wenn auch in den meisten Fällen die mit dem EM Algorithmus ermittelte Kovarianz-Matrix zu gleichen Parameter-Schätzern kommen sollte wie bei Verwendung der FIML-Methode. Enders und Peugh (2004) kommen zu dem Schluss, dass die Standardfehler, die auf einer mit dem EM Algorithmus ermittelten Kovarianz-Matrix basieren, keine Verzerrung aufweisen.

Zur Analyse der Daten der vorliegenden Untersuchung wird auf die Programme SPSS 11.5 und AMOS 5 zurückgegriffen. In SPSS 11.5 ist in einem zusätzlichen Modul der EM Algorithmus implementiert, AMOS 5.0 berücksichtigt FIML. Alle primären Analysen (explorative Faktorenanalysen, Itemanalysen, Reliabilitätsanalysen) werden mit SPSS 11.5 durchgeführt, aus diesem Grund werden fehlende Werte mit EM-Schätzern ersetzt¹⁵, auf denen auch die Kovarianz-Matrix, die als Input für die CFAs und SEMs dient, basiert. Um mögliche Verzerrungen aufzudecken, werden alle finalen Ergebnisse resultierend aus den CFAs und SEMs auch mit FIML berechnet. Dieses Vorgehen sollte gewährleisten, dass keine falschen Schlüsse auf Basis der EM Kovarianz-Matrix gezogen werden, die durch verzerrte Standardfehler verursacht sein könnten.

6.2.3.3 Prüfung auf Normalverteilung und multivariate Ausreißer

Strukturgleichungsanalysen (wie auch alle weiteren auf dem allgemeinen linearen Modell basierenden Analyseverfahren) setzen voraus, dass die Daten multivariat normalverteilt sind. Nicht normalverteilte Daten können bei Verwendung von Standardschätzverfahren wie der

¹⁵ Little (1988) hat einen Test zur Prüfung der Daten auf MCAR entwickelt, der auch in SPSS 11.5 zur Verfügung steht. Sowohl in der Kalibrierungsstichprobe ($\chi^2 = 7994.28$, $df = 7593$, $p < .001$) als auch in der Validierungsstichprobe ($\chi^2 = 7347.5$, $df = 6913$, $p < .001$) kann MCAR angenommen werden. Da auch bei MCAR die konventionellen Verfahren zum Umgang mit fehlenden Werten ineffizient sind, wird trotzdem auf ML-Methoden zurückgegriffen.

Maximum-Likelihood-Methode zu verfälschten Ergebnissen führen¹⁶, West, Finch und Curran (1995) kommen allerdings zu dem Schluss, dass bei Stichprobengrößen von $N = 200 - 500$ und einer Abweichung von der Normalverteilung, die kleiner als Schiefe = 2 und Kurtosis = 7 ist, die ML-Schätzung zu keinen verzerrten Ergebnissen führt (s. a. Muthén & Kaplan, 1985). Weiterhin empfehlen die Autoren Itemparcels¹⁷ zu verwenden. Um diese Voraussetzung für die vorliegende Untersuchung zu prüfen, wurden die zentralen Modellvariablen (Neurotizismus, Ereignisseinschätzung, Ressourcenbewertung, Coping, psychisches Wohlbefinden, körperliche Beschwerden) sowie die Belastungsitems auf Normalverteilung geprüft. Darüber hinaus wurden die Daten auf multivariate Ausreißer hin geprüft, da Ausreißer ebenfalls zu verzerrten Ergebnissen führen können (Tabachnik & Fidell, 2001).

Die Normalverteilungs- und Ausreißeranalyse erfolgte für die Kalibrierungs- und Validierungsstichprobe getrennt mit AMOS 5.0. In Tabelle 9 sind die Ergebnisse der Normalverteilungsanalyse für beide Teilstichproben aufgeführt.

Tabelle 9. Mittelwert und in Klammern Spannweite der univariaten Schiefe und Kurtosis in den zentralen Modellvariablen für die Kalibrierungs- und Validierungsstichprobe

	Kalibrierungsstichprobe ($N = 318$)	Validierungsstichprobe ($N = 318$)
Schiefe	.01 (-2.02 – 2.05)	.02 (-1.85 – 1.98)
Kurtosis	-.14 (-1.31 – 4.29)	-.17 (-1.37 – 3.29)

Die Ergebnisse weisen auf annähernd normalverteilte Daten hin, ausgehend von den Befunden von West et al. (1995) ist davon auszugehen, dass die ML-Methode zu keinen verzerrten Ergebnissen führen wird.¹⁸ Zudem können Hoogland and Boomsma (1998) zeigen, dass die ML-Methode robust gegen kleinere Verletzungen der Normalität ist. Bei Verletzung der Normalverteilungsannahme zeigt darüber hinaus eine Vergleichsstudie von Olsson, Foss, Troye und Howell (2000), das ML-Schätzer im Vergleich zu alternativen Schätzverfahren wie generalized least squares (GLS) oder weighted least squares (WLS) stabiler sind und zu

¹⁶ Folgende Probleme können bei nicht normalverteilten Daten auftreten (West et al., 1995): (a) Der χ^2 -Wert wird mit Zunahme der Abweichung von der Normalverteilung extrem groß, (b) bei kleinen Stichprobengrößen sind ML-basierte χ^2 -Werte (auch bei gegebener Normalverteilung) erhöht, (c) bestimmte Fit-Indizes wie z.B. der CFI werden unterschätzt und (d) Abweichungen von der Normalverteilung können zu geringen Standardfehlern führen, was zu statistisch signifikanten Parametern führen kann, die in der Population aber keine Bedeutung aufweisen.

¹⁷ Mehrere Items werden zu einem Indikator für einen latenten Faktor zusammengefasst (s. weiter unten).

¹⁸ Die Prüfung auf multivariate Kurtosis nach Mardia (1970) zeigt allerdings für beide Teilstichproben eine signifikante Abweichung von der multivariaten Normalverteilung (Kalibrierungsstichprobe: Mardia-Koeffizient = 309.97, $p < .001$; Validierungsstichprobe: Mardia-Koeffizient = 373.15, $p < .001$).

weniger verzerrten Standardfehlern und Fit-Indizes führen (s.a. Chou, Bentler & Satorra, 1991). Um dennoch mögliche Verzerrungen durch die vorliegende Abweichung von der Normalverteilung zu prüfen, werden alle finiten Ergebnisse mit der so genannten Bootstrap-Methode gegen geprüft (Efron, 1979, 1982), die nicht auf der Annahme multivariat normalverteilter Daten beruht.¹⁹ Das prinzipielle Vorgehen bei dieser Methode ist, dass die Stichprobe in eine Vielzahl von Substichproben unterteilt (Zufallsauswahl mit Zurücklegen) und für jede dieser Substichprobe die jeweilige Analyse wiederholt wird. Hierdurch kann die Stabilität von Parametern und von Fit-Indizes geschätzt werden (Byrne, 2001). Die Stichprobengröße sollte bei mindestens $N = 200$ liegen, bei kleineren Stichproben führt die Bootstrap-Methode zu größeren Verzerrungen als ML (Nevitt & Hancock, 2001). Auch die Bootstrap-Methode hat Beschränkungen, eine Übersicht hierüber findet man bei Byrne (2001) und Yung und Bentler (1996). In Amos 5.0 (Arbuckle & Wothke, 1999) ist die Bootstrap-Methode implementiert. Für die vorliegende Untersuchung sind die durch die Bootstrap-Methode ermittelten Standardfehler für alle geschätzten Parameter, die mittleren Parameterschätzer aller Bootstrap-Stichproben und die Differenz (Bias) zwischen den mittleren Bootstrap-Schätzern und den Original-Schätzern (z.B. mit der ML-Methode) relevant. Darüber hinaus erhält man noch die bias-korrigierten Konfidenzintervalle (10%; Efron & Tibshirani, 1993) für die unstandardisierten und standardisierten Parameter. Sollten diese Konfidenzintervalle die Null einschließen, so kann die Hypothese, dass der entsprechende Parameter in der Population Null ist, nicht ausgeschlossen werden.

Die Analyse auf multivariate Ausreißer zeigt, dass sich in der Kalibrierungsstichprobe kein multivariater Ausreißer befindet. In der Validierungsstichprobe muss ein multivariater Ausreißer von den weiteren Analysen ausgeschlossen werden (Bollen, 1987)²⁰, so dass sich die Validierungsstichprobe nach diesem Analyseschritt aus $N = 317$ ReferendarInnen zusammensetzt.

6.2.4 Skalenüberprüfung

Die Skalenprüfung erfolgt ebenfalls in mehreren Schritten: Zunächst werden exploratorische Faktorenanalysen (EFA) für alle Skalen berechnet. Anschließend werden die so ermittelten Skalen mit konfirmatorischen Faktorenanalysen (CFA) geprüft und schließlich die so opti-

¹⁹ Eine weitere mögliche Methode wäre die asymptotically distribution free (ADF) Schätzmethode (Browne, 1984), die allerdings nur bei sehr großen Stichproben ($N \geq 5000$) zu konvergierenden und richtigen Lösungen führt (Chou et al., 1991; Curran, West & Finch, 1996; Hu, Bentler & Kano, 1992) und aus diesem Grund für die vorliegende Untersuchung nicht in Frage kommt.

²⁰ Die Mahalanobis-Distanz diente als Entscheidungskriterium für Ausreißer (Tabachnik & Fidell, 2001).

mierten Skalen auf ihre interne Konsistenz hin analysiert. Alle Analysen erfolgen nur an der Kalibrierungsstichprobe. Erst das Messmodell und schließlich das in den SEMs ermittelte Endmodell werden auch in der Validierungsstichprobe berechnet und gegen geprüft.

6.2.4.1 Explorative Faktorenanalysen

Die exploratorischen Faktorenanalysen dienen der Überprüfung der in anderen Kontexten bewährten Skalen auf Eindimensionalität und auf Dimensionalität der neu entwickelten Skalen. Die gewählte Extraktionsmethode ist ML, da dies auch die gewählte Schätzmethode für alle CFAs und SEMs ist. Darüber hinaus wird bei Extraktion von mehreren Faktoren schiefwinklig rotiert (PROMAX; s. Tabachnik & Fidell, 2001), da bei vorliegender Mehrdimensionalität ein Zusammenhang zwischen den ermittelten Faktoren anzunehmen ist. Es werden nur solche Faktorenladungen zugelassen, die über $a = .35$ liegen. Nebenladungen sollten kleiner $a = .20$ sein, wobei auch das Verhältnis zur Hauptladung in Betracht gezogen wird (Rost & Schermer, 1986). Diese Grenzen für Haupt- und Nebenladungen werden gewählt, da bei der gewählten Rotationsmethode niedrige Ladungen (nach Varimax-Rotation) größer und höhere Ladungen niedriger werden (Tabachnik & Fidell, 2001).

Belastungsskala. Der Scree-Plot deutet auf eine vier-dimensionale Struktur der 23 Belastungsitems hin (Eigenwertverlauf: 4.7, 2.17, 1.88, 1.63, 1.23, 1.06, 1.04, 0.94, etc.). In Tabelle 10 ist das rotierte Ladungsmuster aufgeführt.

Tabelle 10. Ladungsmuster (nach Rotation) der Items der Belastungsskala auf den vier extrahierten Faktoren (Faktorladungen von $a < .20$ sind nicht aufgeführt; Items mit fett markierter Ladung werden ausgewählt)

Item ¹	Faktor 1	Faktor 2	Faktor 3	Faktor 4
belast01	.26			
belast02	.93			
belast03	.72			
belast04	.30	.36		-.21
belast05		.62		
belast06		.46		
belast07	.22	.31		
belast08	.28			.24
belast09				.66
belast10	.69			
belast11		.49		
belast12		.34		.22
belast13		.37		
belast14				.61
belast15		.59		
belast16			-.20	.48
belast17		.41		
belast18			.94	
belast19			.88	
belast20			.24	
belast21				.32
belast22		.27		
belast23			.27	.62

Anmerkung: 1 Für den genauen Itemwortlaut s. Tabelle 5 (S. 79)

Die vier Faktoren lassen sich inhaltlich gut benennen und interpretieren. *Faktor 1*, gebildet aus den Items belast02 (mangelnde Motivation bzw. Konzentrationsfähigkeit der SchülerInnen), belast03 (Disziplinprobleme) und belast10 (Geringe Lernbereitschaft von SchülerInnen), umfasst Belastungen durch Schülerinnen und Schüler, der Faktor erklärt 20.43% der Varianz. *Faktor 2* beinhaltet die Items belast05 (Ärger mit Behörden bzw. Institutionen), belast06 (Probleme mit den Eltern), belast11 (Verwaltungsarbeit), belast13 (hektisches und störungsvolles Arbeitsklima an der Schule), belast15 (Konflikte mit der Ausbildungsschule) und belast17 (Konflikte mit AusbilderInnen) und vereint Belastungen durch Konflikte in verschiedenen Bereichen, aber auch durch z.B. Verwaltungsarbeit. Insgesamt bindet dieser Faktor 9.43% der Varianz. Auf *Faktor 3* laden lediglich die Items belast18 (Pädagogische Prüfungsarbeit) und belast19 (Examensprüfungen) substantiell, beide Items beziehen sich auf die zukünftigen Prüfungen (Pädagogische Prüfungsarbeit und Examensprüfungen), die Varianzaufklärung durch diesen Faktor beträgt 8.17%. Schließlich umfasst *Faktor 4* mit den Items belast09 (Vor- und Nachbereitung des Unterrichts), belast14 (Unterrichtsproben/Lehrproben), belast16 (eigenverantwortlicher Unterricht), belast23 (Unterrichtsentwürfe) Belastungen, die mit dem eigenen Unterricht zusammenhängen, die Varianzaufklärung durch diesen Faktor beträgt 7.08%. Alle übrigen Items der Belastungsskala lassen sich keinem der Faktoren eindeutig zuordnen und werden somit in allen weiteren Analysen nicht weiter berücksichtigt. Die Faktoren korrelieren moderat miteinander (s. Tabelle 11).

Tabelle 11. Interkorrelation der vier Belastungsfaktoren

	Faktor 1	Faktor 2	Faktor 3	Faktor 4
Faktor 1	1	.27***	.09	.29***
Faktor 2		1	.25***	.27***
Faktor 3			1	.22***
Faktor 4				1

Anmerkung: *** $p < .001$

Neurotizismus. Wie erwartet laden alle Items der Neurotizismusskala ausreichend hoch auf dem ersten Faktor (Minimum: $a = .47$), der Scree-Plot weist deutlich auf eine eindimensionale Struktur hin (Eigenwertverlauf: 5.48, 0.95, 0.89, etc.), der erste Faktor bindet 45.67% der Varianz.

Ausmaß an sozialer Unterstützung. Wie bereits weiter oben begründet, werden zwei Items der Skala zur sozialen Unterstützung nicht weiter berücksichtigt, weil sie von einem substanziellen Teil der Befragten nicht beantwortet werden konnten (fehlendes MentorInnen-system). Somit umfasst die Skala 10 Items. Eine EFA über alle 10 Items führt zu Problemen bei der Kommunalitätsschätzung bei einzelnen Items, so dass die Ergebnisse dieser Analyse nur schwer interpretierbar sind. Eine mögliche Ursache kann die tabellarische Vorgabeform der Skala sein, die möglicherweise bestimmte Antwortmuster (spaltenweise vs. zeilenweise Beantwortung der Zellen) provoziert. Als Lösung des Problems wird jeweils eine EFA für die Items der instrumentellen Unterstützung und der Items der emotionalen Unterstützung berechnet. In beiden Fällen ergibt sich eine eindeutig eindimensionale Lösung: Der erste Faktor für die Items der instrumentellen Unterstützungsskala klärt 37.39% der Varianz auf (Eigenwertverlauf: 1.87, 0.98, etc.), die geringste Ladung liegt bei $a = .35$. Die Varianzaufklärung für den ersten Faktor der emotionalen Unterstützungsskala liegt bei 38.52% (Eigenwertverlauf: 1.93, 0.88, etc.), die geringste Ladung liegt hier bei $a = .37$.

Ereigniseinschätzung. Die Ereigniseinschätzung umfasst, wie im Methodenteil beschrieben, Items zur Bedrohungs-, Verlust- und Herausforderungsbewertung. Demnach wäre eine dreidimensionale Struktur zu erwarten. Die EFA legt aber nahe, dass den Items nur ein gemeinsamer Faktor unterliegt (Eigenwertverlauf: 3.77, 1.05, .76, etc.; Varianzaufklärung des ersten Faktors: 47.11%), die geringste Itemladung liegt bei $a = .45$. Die beiden Herausforderungsitems laden negativ auf diesem ersten Faktor, so dass dieser Faktor als bipolare Bewertungsdimension mit den Endpolen „Herausforderungsbewertung“ auf der einen Seite und „Bedrohungs- und Verlustbewertung“ auf der anderen Seite interpretiert werden kann.

Bei Vorgabe von drei Faktoren lädt lediglich ein Herausforderungsitem auf einem eigenen Faktor, das andere lädt hoch auf einem zweiten Faktor. Auch die übrigen Items weisen zum Teil substanzielle Nebenladungen auf.

Selbstwirksamkeitserwartungen. Für die vorliegende Untersuchung wurde, wie im Methodenteil dargelegt, auf Items zur Messung der allgemeinen und der situationsspezifischen Selbstwirksamkeitserwartungen zurückgegriffen. Zu erwarten wäre, dass die ausgewählten Selbstwirksamkeitsitems auf zwei positiv miteinander korrelierten Faktoren laden müssten. Eine EFA mit allen Items zeigt allerdings, dass sich eine zweifaktorielle Lösung nicht anbietet: Der Eigenwertverlauf (3.71, 1.07, 0.95, etc.) spricht deutlich für eine einfaktorielle Lösung, die Varianzaufklärung durch den ersten Faktor liegt bei 37.2%. Die Vorgabe von zwei Faktoren zeigt, dass auf einem zweiten Faktor lediglich ein Item substanziell lädt. Da aber im vorliegenden Untersuchungskontext situationspezifische Selbstwirksamkeitser-

wartungen als Indikator für die Ressourceneinschätzung dienen sollen und eine inhaltliche Trennung der Items zu allgemeinen und situationsspezifischen Selbstwirksamkeitserwartungen wichtig ist (Schmitz & Schwarzer, 2000), werden nur die Items zu berufsbezogenen Selbstwirksamkeitserwartungen herangezogen. Eine EFA mit diesen fünf Items weist auf eine eindeutig eindimensionale Lösung hin (Eigenwertverlauf: 1.99, 0.90, etc.), der erste Faktor bindet 39.73% der Varianz. Allerdings weist ein Item nur eine Faktorladung von $a < .29$ auf. Um die inhaltliche Breite und die Itemanzahl nicht zu stark zu reduzieren, wird dieses Item zunächst in der Skala belassen. Alle übrigen Items laden ausreichend hoch auf diesem Faktor (geringste Ladung: $a = .43$). Der Faktor bildet einen Indikator für situationsspezifische Selbstwirksamkeitserwartungen.

Coping. Die EFA mit allen Coping-Items führt bei der Extraktion von Faktoren zu Problemen bei den Kommunalitätsschätzungen, so dass auf die Prüfung der Skala auf Mehrdimensionalität verzichtet wird. Vielmehr wird in Anlehnung an van Dick (1999a) eine eindimensionale Lösung mit zwei Endpolen (emotionszentriertem vs. problemzentriertem Coping) vorgegeben²¹. Bei Vorgabe von einem Faktor kommt es bei der Extraktion zu keinen Problemen bei den Kommunalitätsschätzungen. Der erste Faktor klärt allerdings nur 22.3% der Varianz auf, der Eigenwertverlauf spricht aber für die eindimensionale Lösung (3.57, 1.87, 1.47, 1.29, 1.03, 0.89, etc.). Allerdings werden vier Items wegen zu geringer Ladung ($a < .30$) auf dem Faktor nicht in die Skala aufgenommen. Die Skala umfasst somit nach diesem Analyseschritt 11 Items, wovon 6 Items problemzentriertes und 5 Items emotionszentriertes Bewältigungsverhalten messen. Im Anhang ist das genaue Ladungsmuster aufgeführt (S. XXII).

Körperliche Beschwerden. Erwartet wird für die Beschwerdeskala eine eindimensionale Lösung, die sich so auch in einer EFA zeigt (Eigenwertverlauf: 2.85, 1.17, 0.98, etc.). Der erste Faktor bindet 31.69% der Varianz. Zwei Items der Skala laden allerdings nur schwach auf dem ersten Faktor ($a < .30$) und werden aus diesem Grund ausgeschlossen (es handelt sich um die Beschwerden „Sodbrennen und saures Aufstoßen“ und „Stimmprobleme“), alle übrigen sieben Items weisen eine ausreichend hohe Faktorladung auf ($a > .42$).

Psychisches Wohlbefinden. Wie bei der Copingskala ergeben sich auch hier bei der EFA Probleme bei Kommunalitätsschätzungen. Da eine eindimensionale Struktur erwartet wird, wurde ein Faktor vorgegeben. Hier kommt es zu keinen Problemen bei den Kommunalitätsschätzungen. Der Faktor klärt 50.85% der Varianz auf (Eigenwertverlauf: 4.58, 1.18, 0.89, etc.), alle Items weisen eine ausreichend hohe Faktorladung auf (geringste Faktorladung: $a = .47$).

²¹ Die Vorgabe in SPSS von einem emotions- und einem problemzentriertem Faktor führt ebenfalls zu keiner befriedigenden Lösung, da eine Reihe von Items nicht auf ihrem inhaltlich entsprechendem Faktor laden.

6.2.4.2 Konfirmatorische Faktorenanalysen

Im Folgenden werden die Ergebnisse der CFAs für jede Skala vorgestellt. Zum Teil mussten Modifikationen vorgenommen werden, auf die im Folgenden eingegangen wird. Bei den Modifikationen wurden auch Fehlerkorrelationen berücksichtigt. Hierbei wurde aber nur nach inhaltlichen Gesichtspunkten vorgegangen (z.B. semantische Ähnlichkeit, Antwortformat wie rekodierte Variablen), um eine rein datengesteuerte Modellanpassung zu vermeiden (MacCallum, Roznowski & Necowitz, 1992).

Belastungsskala. Eine Prüfung der Belastungsskala mit den in der exploratorischen Faktorenanalyse ermittelten vier korrelierten Subdimensionen weist eine unbefriedigende Datenanpassung auf ($\chi^2 = 274.24$; $df = 80$; $p < .001$; $CFI = .86$; $RMSEA = .09$; $SRMR = .07$; $\chi^2/df = 3.27$). Zur Verbesserung des Fits muss eine Fehlerkorrelation zwischen zwei Items der Subskala „Belastungen durch Konflikte und Verwaltungsarbeit“ zugelassen werden sowie eine Fehlervarianz auf Null restringiert werden. Der Modell-Fit ist allerdings auch nach diesen Modifikationen nach gängigen Kriterien nicht ausreichend ($\chi^2 = 235.96$; $df = 83$; $p < .001$; $CFI = .89$; $RMSEA = .08$; $SRMR = .07$; $\chi^2/df = 2.84$), wenn auch besser als das Ausgangsmodell ($\Delta\chi^2 = 38.28$, $df = 3$, $p < .01$). Eine eindimensionale Lösung ($\chi^2 = 909.89$; $df = 90$; $p < .001$; $CFI = .39$; $RMSEA = .17$; $SRMR = .13$; $\chi^2/df = 10.11$) zeigt zusätzlich eine signifikant schlechtere Anpassung an die Daten auf ($\Delta\chi^2 = 673.93$, $df = 7$, $p < .01$).

Neurotizismus. Auch hier musste eine Fehlerkorrelation zugelassen werden zwischen zwei Items („Ich fühle mich selten einsam oder traurig“ und „Ich bin selten traurig oder deprimiert“, beide Items sind semantisch sehr ähnlich und haben einen sehr ähnlichen Satzaufbau). Die Passung dieses modifizierten Modells ist als gut zu bezeichnen ($\chi^2 = 116.729$; $df = 53$; $p < .001$; $CFI = .96$; $RMSEA = .06$; $SRMR = .04$; $\chi^2/df = 2.20$).

Soziale Unterstützung. Zunächst wurden die beiden Subskalen instrumentelle Unterstützung und praktische Unterstützung, wie in der exploratorischen Faktorenanalyse, separat überprüft. Beide Subskalen weisen eine ausreichende Passung auf (Instrumentelle Unterstützung: $\chi^2 = 12.18$; $df = 5$; $p < .05$; $CFI = .94$; $RMSEA = .07$; $SRMR = .04$; $\chi^2/df = 2.44$; Emotionale Unterstützung: $\chi^2 = 5.44$; $df = 5$; $p = .37$; $CFI = 1.00$; $RMSEA = .02$; $SRMR = .03$; $\chi^2/df = 1.09$). Anschließend wurde ein Zwei-korrelierte-Faktoren-Modell geprüft. Dieses Modell weist eine schlechte Datenanpassung auf ($\chi^2 = 536.58$; $df = 34$; $p < .001$; $CFI = .45$; $RMSEA = .22$; $SRMR = .13$; $\chi^2/df = 15.78$). Werden aber die Korrelationen zwischen den Fehlertermen der Items einer Unterstützungsquelle zugelassen, so verbessert sich der Fit deutlich ($\Delta\chi^2 = 498.01$, $df = 5$, $p < .01$) und kann als sehr gut bezeichnet werden ($\chi^2 = 38.57$;

df = 29; $p = .11$; CFI = .99; RMSEA = .03; SRMR = .04; $\chi^2/\text{df} = 1.33$). Allerdings korrelieren beide Faktoren mit $r = .69$ sehr hoch miteinander, was eine getrennte Betrachtung beider Faktoren als ungeeignet erscheinen lässt. Eine eindimensionale Lösung weist ebenfalls nur einen schlechten Fit auf ($\chi^2 = 538.24$; df = 35; $p < .001$; CFI = .45; RMSEA = .21; SRMR = .13; $\chi^2/\text{df} = 15.38$). Zwar lässt sich das eindimensionale Modell signifikant verbessern (Korrelation der Fehlerterme der Items einer Unterstützungsquelle; $\Delta\chi^2 = 426.2$, df = 5, $p < .001$), die Modellanpassung ist aber immer noch unbefriedigend ($\chi^2 = 112.04$; df = 30; $p < .001$; CFI = .91; RMSEA = .09; SRMR = .05; $\chi^2/\text{df} = 15.38$) und die Faktorladungen liegen allesamt in einem niedrigen Bereich ($a_{\min} = .29$, $a_{\max} = .58$). Somit sind sowohl die ein- als auch die zweidimensionale Lösung nicht optimal, auf dieses Problem wird an späterer Stelle näher eingegangen werden. Zunächst wird eine eindimensionale Lösung präferiert.

Ereigniseinschätzung. Die eindimensionale Bewertungsskala mit den beiden Endpolen „Herausforderung“ und „Verlust/Bedrohung“ zeigt eine schlechte Datenanpassung ($\chi^2 = 154.72$; df = 20; $p < .001$; CFI = .84; RMSEA = .15; SRMR = .07; $\chi^2/\text{df} = 7.74$). Wird die Fehlerkorrelation zwischen zwei Verlust-Items zugelassen, so verbessert sich der Fit deutlich und signifikant ($\Delta\chi^2 = 78.16$, df = 1, $p < .001$) und kann als knapp befriedigend bezeichnet werden ($\chi^2 = 76.56$; df = 19; $p < .001$; CFI = .93; RMSEA = .10; SRMR = .06; $\chi^2/\text{df} = 4.03$). Eine dreidimensionale Lösung zeigt eine gute Datenanpassung ($\chi^2 = 58.25$; df = 17; $p < .001$; CFI = .95; RMSEA = .09; SRMR = .05; $\chi^2/\text{df} = 3.43$), allerdings sind die Korrelationen zwischen den latenten Faktoren sehr hoch ($r_{\min} = .65$, $r_{\max} = .95$), so dass eine Trennung nicht sinnvoll erscheint (vgl. Scheck & Kinicki, 2000).

Selbstwirksamkeitserwartungen. Wie oben begründet, werden nur die fünf berufsbezogenen Items verwendet. Die Skala zeigt eine gute Datenanpassung ($\chi^2 = 8.54$; df = 5; $p = .13$; CFI = .98; RMSEA = .05; SRMR = .03; $\chi^2/\text{df} = 1.71$), wenn auch ein Item eine relativ niedrige Faktorladung aufweist ($a = .29$).

Coping. Die eindimensionale Copingskala zeigt eine unbefriedigende Datenanpassung ($\chi^2 = 387.90$; df = 44; $p < .001$; CFI = .58; RMSEA = .16; SRMR = .10; $\chi^2/\text{df} = 8.82$). In mehreren Modifikationsschritten werden insgesamt fünf Fehlerkorrelationen zugelassen sowie ein Item aufgrund einer zu geringen Faktorladung von $a = .17$ ausgeschlossen, das so modifizierte Modell weist eine gute ($\chi^2 = 66.86$; df = 30; $p < .001$; CFI = .95; RMSEA = .06; SRMR = .05; $\chi^2/\text{df} = 2.23$) und signifikant bessere Datenanpassung auf als das Ausgangsmodell ($\Delta\chi^2 = 357.93$, df = 14, $p < .001$). Fehlerkorrelationen wurden zwischen vier emotions-

zentrierten sowie zwischen zwei problemzentrierten Copingitems zugelassen, die sich beide auf die Suche nach sozialer Unterstützung beziehen.

Beschwerden. Auch hier zeigt das eindimensionale Ausgangsmodell eine schlechte Datenanpassung ($\chi^2 = 67.70$; $df = 14$; $p < .001$; CFI = .86; RMSEA = .11; SRMR = .06; $\chi^2/df = 4.84$). Es werden zwei Fehlerkorrelationen zugelassen, diese Modifikation führt zu einer deutlichen Verbesserung des Fits ($\Delta\chi^2 = 29.97$, $df = 2$, $p < .001$), die Modellanpassung kann nun als ausreichend angesehen werden ($\chi^2 = 37.73$; $df = 12$; $p < .001$; CFI = .93; RMSEA = .08; SRMR = .05; $\chi^2/df = 3.14$).

Psychisches Wohlbefinden. Das eindimensionale Modell hat eine schlechte Anpassung an die Daten ($\chi^2 = 205.32$, $df = 14$; $p < .001$; CFI = .83; RMSEA = .21; SRMR = .07; $\chi^2/df = 14.67$). Nachdem zwei Fehlerkorrelationen zwischen zwei inhaltlich sehr ähnlichen Indikatoren zugelassen wurden, zeigt das so modifizierte Modell eine signifikant bessere Datenanpassung ($\Delta\chi^2 = 158.80$, $df = 2$, $p < .001$), der Fit kann als befriedigend bezeichnet werden ($\chi^2 = 46.52$; $df = 12$; $p < .001$; CFI = .97; RMSEA = .10; SRMR = .05; $\chi^2/df = 3.88$).

Die CFA für alle Skalen separat zeigen, dass bis auf eine Ausnahme (Selbstwirksamkeitsskala) Modifikationen vorgenommen werden mussten, um zumindest annähernd ausreichende Modellanpassungen zu bekommen. Dies weist auf Probleme auf der Messebene hin. Gerade im Hinblick auf Strukturgleichungsanalysen kann dies zu Problemen führen insofern, dass ein möglicherweise passendes Modell abgelehnt wird. Um die Probleme auf der Messebene abfangen zu können, wird auf die Methode des Itemparcelings²² zurückgegriffen (Little, Cunningham, Shahar & Widaman, 2002). Parcels können definiert werden als ein Indikator auf der Aggregatebene, der sich aus der Summe (oder dem Mittel) von zwei oder mehr Items zusammensetzt. Es gibt eine Reihe von Argumenten für Parcels, die sowohl theoretischer als auch empirischer Art sind. Für die theoretische Argumentation sei auf Little et al. (2002) verwiesen. Wichtiger für die vorliegende Studie sind empirische Argumente für (und gegen) Parceling. Im Vergleich zu Daten auf der aggregierten Ebene (hier auf der „Parcel-Ebene) haben Daten auf der Itemebene eines oder mehrere der folgenden Nachteile: (a) geringere Reliabilität, (b) geringere Kommunalität, (c) einen geringeren Quotienten zwischen gemeinsamer und eigener Faktorvarianz sowie (d) eine größere Wahrscheinlichkeit für Verteilungsabweichungen (Little et al., 2002). Darüber hinaus haben Items geringere, größere oder ungleichere Intervalle zwischen Skalenpunkten als Parcels (Bagozzi & Heatherton, 1994; Kishton & Widaman, 1994; McCallum, Widaman, Zhang & Hong, 1999). Des Weiteren argumentieren

²² Für Parceling gibt es keine direkt passende Übersetzung, aus diesem Grund wird der englische Originalbegriff beibehalten.

Befürworter von Itemparceling, dass gerade bei kleineren Stichproben Parcels Items vorgezogen werden sollten, da weniger Parameter zur Definition eines Konstrukts benötigt werden und somit der Modell-Fit besser wird (Bagozzi & Edwards, 1998; Bagozzi & Heatherton, 1994). Modelle, die auf Parcels basieren, sind im Vergleich zu Modellen mit Daten auf der Itemebene, (a) sparsamer (weniger geschätzte Parameter, sowohl lokal um ein Konstrukt zu definieren, als auch global in der Repräsentation eines Modells), (b) die Wahrscheinlichkeit für Fehlerkorrelationen und Nebenladungen werden reduziert (zum einen, weil weniger Indikatoren verwendet werden und zum anderen, weil die spezifischen Varianzen kleiner sind) und c) führen zu einer Verringerung von verschiedenen Quellen falscher Stichprobenziehungen. Auf der anderen Seite gibt es einige Dinge, die beim Parceling beachtet werden müssen. So wird von einigen Autoren darauf hingewiesen, dass das zugrunde liegende Konstrukt eindimensional sein sollte, nur dann sei Parceling indiziert (Bagozzi & Heatherton, 1994; Bandalos, 2002; Bandalos & Finney, 2002). Ein weiteres Problem ist, dass durch Parceling wichtige Informationen verloren gehen können. Allerdings ist dieses Problem abhängig von der Fragestellung (Bagozzi & Heatherton, 1994). Auch Little et al. (2002) argumentieren in diese Richtung: Wenn Ziel einer Untersuchung ist, die Zusammenhänge zwischen einzelnen Items zu verstehen (wie z.B. die Entwicklung einer neuen Skala), so sollten keine Parcels verwendet werden. Wenn dagegen Ziel ist, die Zusammenhänge zwischen latenten Konstrukten zu untersuchen, so ist Parceling indiziert, da hier von Wichtigkeit ist, reine latente Faktoren zu erhalten (ohne Nebenladungen von Items oder Fehlerkorrelationen).

Für die vorliegende Untersuchung scheint die Parcel-Methode indiziert zu sein, um die Probleme auf der Messebene in den Griff zu bekommen. Primäres Ziel der Untersuchung ist die Überprüfung der transaktionalen Stresstheorie, d.h. die Beziehung zwischen verschiedenen latenten Konstrukten. Von geringerem bzw. keinem Interesse sind die genauen Beziehungen zwischen den Items untereinander. Gegen Parceling könnte sprechen, dass einige der Konstrukte wie die Ereignisseinschätzung aus theoretischer Perspektive mehrdimensional sind. Die EFAs und CFAs zeigen aber, dass dies für keines der Konstrukte bzw. Skalen der Fall ist. Um aber ausgewogene Parcels zu generieren, wird bei (theoretisch) mehrdimensionalen Konstrukten auf die Methode von Kishton und Widaman (1994) zurückgegriffen²³, die die Mehrdimensionalität von Items berücksichtigt. Dabei werden Items zu einem Parcel zusammengefasst, die jeweils verschiedene Facetten eines Konstrukts abbilden, so dass jedes Parcel wiederum alle Facetten beinhaltet. Dies würde z.B. im vorliegenden Falle für die Bewertungsskala bedeuten, dass in jedem Parcel alle drei theoretischen Facetten (Herausforderung,

²³ Einen Überblick über weitere Methoden zur Bildung von Parcel geben Little et al. (2002).

Verlust, Bedrohung) idealerweise enthalten sein sollten. Kishton und Widaman (1994) können zeigen, dass diese Methode (bei gegebener Mehrdimensionalität) Vorteile gegenüber einer Methode hat, die Parcel nach der internen Konsistenz bildet. Die Methode von Kishton und Widaman wird bei der Bewertungs- und Copingskala verwendet. Bei allen anderen Skalen wird auf eine Methode zurückgegriffen, die die Parcels eines Konstruktes ausbalanciert. Dabei dienen die Faktorladungen der einzelnen Items als Maßstab: Zunächst werden (wenn drei Parcels gebildet werden sollen) die drei Items mit den höchsten Ladungen als Anker für die Parcels ausgewählt. Im nächsten Schritt werden die drei Items mit den nächst höheren Ladungen ausgewählt und in umgekehrter Reihenfolge den Parcels zugeordnet, so dass dem Parcel mit dem Item mit der höchsten Ladung nun das Item zugeordnet wird, welches von den drei nächsten Items die niedrigste Ladung hat. Unter manchen Bedingungen können Parcels eine verschiedene Anzahl von Items haben, um eine ausreichende Balance zu erhalten (s. Little et al., 2002).

Für die vorliegende Untersuchung wurden für jedes Konstrukt drei Parcels nach den oben beschriebenen Methoden mit Ausnahme der Skala zur Sozialen Unterstützung, für die vier Parcels gebildet wurden, generiert. Die genaue Item-zu-Parcel-Zuordnung kann dem Anhang entnommen werden (S. XXV). Auf das Parceling für die Skala zur sozialen Unterstützung soll an dieser Stelle etwas näher eingegangen werden, da hier von dem beschriebenen Vorgehen leicht abgewichen wurde. Wie weiter oben beschrieben, korrelieren bei Vorgabe von einem latenten Faktor die beiden Subdimensionen „instrumentelle Unterstützung“ und „emotionale Unterstützung“ relativ hoch miteinander, zusätzlich kann nur dann ein akzeptabler Fit erzielt werden, wenn die Fehlerkorrelation der Items zu einer Quelle von sozialer Unterstützung zugelassen werden. Eine eindimensionale Lösung weist wiederum eine schlechte Datenanpassung auf, darüber hinaus liegen hier die Faktorladungen der Items in einem niedrigen Bereich. Beide Subdimensionen zeigen in getrennten CFAs eine gute Datenanpassung. Zur Lösung dieses Problems werden auf Basis der getrennten CFAs für jede Subdimension zwei Parcel gebildet, die wiederum zusammen einen gemeinsamen Faktor bilden werden. Diese Lösung zeigt in einer eigenen CFA eine gute Modellanpassung ($\chi^2 = 6.93$, $df = 2$, $p < .001$, $CFI = .98$, $RMSEA = .09$, $SRMR = .03$, $\chi^2/df = 3.48$), so dass davon ausgegangen werden kann, dass das Vorgehen angemessen ist.

Somit sind mit Ausnahme der Skala zur sozialen Unterstützung alle latenten Konstrukte lokal identifiziert, was den Vorteil hat, dass es nur eine Lösung gibt, die den latenten Faktor am besten repräsentiert (Little et al., 2002).

Im nächsten Schritt ist zu prüfen, ob das Messmodell auf Parcel-Ebene eine befriedigende Datenanpassung aufweist und sich in der Validierungsstichprobe ebenfalls bewährt.

6.2.5 Überprüfung des Messmodells

Die Überprüfung des Messmodells erfolgt in zwei Schritten: Zunächst wird an der Kalibrierungsstichprobe geprüft, ob das Messmodell eine ausreichende Anpassung an die Daten aufweist. Anschließend wird das Messmodell an der Validierungsstichprobe überprüft und kreuzvalidiert. Abschließend wird analysiert, ob es zwischen Kalibrierungs- und Validierungsstichprobe Mittelwertsunterschiede in den latenten Konstrukten gibt.

6.2.5.1 Überprüfung des Messmodells an der Kalibrierungs- und Validierungsstichprobe

Zur Überprüfung des Messmodells wurde eine CFA mit sieben korrelierten latenten Faktoren mit jeweils den entsprechenden Parcel-Indikatoren durchgeführt. Das Modell zeigt eine gute Anpassung an die Daten ($\chi^2 = 325.65$, $df = 188$, $p < .001$, $CFI = .96$, $RMSEA = .05$, $p\text{-close} = .63$, $SRMR = .05$, $\chi^2/df = 1.73$).²⁴ In Tabelle 12 sind Faktorladungen der Parcels und die Interkorrelation zwischen den latenten Konstrukten aufgeführt.

²⁴ Auf Basis der standardisierten Residuen wird beurteilt, ob es signifikante Abweichungen zwischen der Kovarianzmatrix der Stichprobe und der eingeschränkten (Modell) Kovarianzmatrix gibt. In der Kalibrierungsstichprobe liegen lediglich 16 von 484 Residuen über dem kritischen Wert von 2 (ca. 3%).

Tabelle 12. Faktorladungen und Interkorrelationen des Messmodells (Kalibrierungsstichprobe)

Model	Goodness of Fit						
Null-Modell	$\chi^2 (231, N = 318) = 3425,17, p < .001, \chi^2/df=14.83$						
7 (korrelierte) Faktoren-Modell	$\chi^2 (188, N = 318) = 325.65, p < .001, CFI = .96, RMSEA = .05, p-close = .63,$ SRMR = .05, $\chi^2/df = 1.73$						
Faktorladungen							
Neurotizismus	Soziale Unterstützung	Ereigniseinschätzung	Selbstwirksamkeits- einschätzung	Coping	Beschwerden	Psychisches Wohlbefinden	
.82	.66	.76	.70	.64	.76	.81	
.83	.68	.90	.58	.68	.61	.81	
.87	.61	.83	.51	.85	.64	.85	
	.67						
Interkorrelationen							
	1	2	3	4	5	6	7
1 Neurotizismus	1	-.36***	.75***	-.51***	-.56***	.65***	-.82***
2 Soziale Unterstützung		1	-.34***	.46***	.49***	-.08	.43***
3 Ereigniseinschätzung			1	-.55***	-.57***	.49***	-.74***
4 Selbstwirksamkeitserwartungen				1	.56***	-.16	.52***
5 Coping					1	-.33***	.61***
6 Beschwerden						1	-.64***
7 Psychisches Wohlbefinden							1

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

In der Validierungsstichprobe zeigt das Modell ebenfalls eine ausreichende, wenn auch etwas schlechtere, Anpassung an die Daten ($\chi^2 = 417.58$, $p < .001$, CFI = .93, RMSEA = .06, p-close = .01, SRMR = .05, $\chi^2/df = 2.22$).²⁵ In Tabelle 13 sind alle Faktorladungen und Interkorrelationen aufgeführt.

²⁵ Auch in der Validierungsstichprobe liegen lediglich 8 von 484 standardisierten Residuen über dem kritischen Wert von 2 (ca. 2 %).

Tabelle 13. Faktorladungen und Interkorrelationen des Messmodells (Validierungsstichprobe)

Modell		Goodness of Fit					
Null-Modell		$\chi^2 (231, N = 317) = 3735.82, p < .001, \chi^2/df = 16.17$					
7 (korrelierte) Faktoren-Modell		$\chi^2 (188, N = 317) = 417.58, p < .001, CFI = .93, RMSEA = .06; p\text{-close} = .01, SRMR = .05, \chi^2/df = 2.22$					
Faktorladungen							
Neurotizismus	Soziale Unterstützung	Ereigniseinschätzung	Selbstwirksamkeitseinschätzung	Coping	Beschwerden	Psychisches Wohlbefinden	
.83***	.76***	.76***	.63***	.78***	.64***	.81***	
.87***	.73***	.89***	.69***	.81***	.68***	.79***	
.80***	.68***	.82***	.61***	.74***	.67***	.84***	
	.63***						
Interkorrelationen							
	1	2	3	4	5	6	7
1 Neurotizismus	1	-.29***	.84***	-.38***	.58***	.68***	-.84***
2 Soziale Unterstützung		1	-.33***	.38***	.38***	-.18*	.39***
3 Ereigniseinschätzung			1	-.56***	-.71***	.43***	-.78***
4 Selbstwirksamkeitserwartungen				1	.67***	-.19*	.35***
5 Coping					1	-.38***	.54***
6 Beschwerden						1	-.70***
7 Psychisches Wohlbefinden							1

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

Da es Hinweise darauf gibt, dass die vorliegenden Daten von einer multivariaten Normalverteilung abweichen (s.o.), wurde das Messmodell sowohl für die Kalibrierungs- als auch für die Validierungsstichprobe mit der Bootstrap-Methode²⁶ geschätzt. Die Ergebnisse der Parameter-Schätzung zeigen in beiden Stichproben, dass die Standardfehler der (unstandardisierten) Faktorladungen der Parcels wie auch die (unstandardisierten) Faktorinterkorrelationen mit der Bootstrap-Methode höher sind als mit der ML-Methode, im Schnitt liegt die Zunahme in der Kalibrierungsstichprobe bei $M = 11.01\%$ ($SD = 17.28$, $Min = -8.47$, $Max = 50.53$) für die Faktorladungen (Validierungsstichprobe: $M = 11.13\%$, $SD = 17.70$, $Min = -13.79$, $Max = 46.83$) und $M = 7.69\%$ ($SD = 14.27$, $Min = -15.87$, $Max = 34.55$) für die Faktorinterkorrelationen (Validierungsstichprobe: $M = 8.27\%$, $SD = 10.82$, $Min = -12.50$, $Max = 25.00$), somit ist die Abweichung in einem tolerablen Bereich. Der Bias in den (unstandardisierten) Parameterschätzern fällt für beide Stichproben ebenfalls nur gering aus, sowohl für die Faktorladungen (Kalibrierungsstichprobe: $M = .009$, $SD = .012$, $Min = -.007$, $Max = .033$; Validierungsstichprobe: $M = .006$, $SD = .007$, $Min = -.005$, $Max = .020$) als auch für die Faktorinterkorrelationen (Kalibrierungsstichprobe: $M = -.001$, $SD = .005$, $Min = -.014$, $Max = .006$; Validierungsstichprobe: $M = .000$, $SD = .003$, $Min = -.005$, $Max = .006$). Darüber hinaus zeigen die bias-korrigierten Konfidenzintervalle, dass alle Faktorladungen größer als Null sind. Auch die bias-korrigierten Konfidenzintervalle für die mit der ML-Methode signifikanten Faktorinterkorrelationen schließen nicht die Null ein. Insgesamt kann resümiert werden, dass die Verzerrungen, die möglicherweise durch die Abweichung der Daten von der multivariaten Normalverteilung verursacht sein können, gering ausfallen und auf die Parameterschätzer nur geringen Einfluss ausüben. Alle Parameter, die unter Verwendung der ML-Methode statistisch signifikant sind, sind dies auch bei Verwendung der Bootstrap-Methode. Die folgenden Analysen werden aus diesem Grund weiter mit der ML-Methode berechnet.

Sowohl für die Kalibrierungs- als auch für die Validierungsstichprobe wurde das Messmodell auch mit einem Datensatz berechnet, in dem die fehlenden Werte nicht durch EM-geschätzte Werte ersetzt wurden, sondern die Modellparameter direkt mit der FIML-Methode (s.o.) berechnet wurden. In beiden Stichproben führt dieses Verfahren zu gleichen Ergebnissen, sowohl in Hinblick auf die Fit-Indizes als auch in Hinblick auf die Modellparameter. Somit kann eine Verzerrung der Ergebnisse verursacht durch unterschiedliche Methoden zum Umgang mit fehlenden Werten ausgeschlossen werden.

²⁶ Es wurden 250 Bootstrap-Stichproben verwendet. Nevitt und Hancock (2001) können zeigen, dass eine noch größere Anzahl von Bootstrap-Stichproben zu keinen besseren Ergebnissen führt. Zur näheren Beschreibung der Bootstrap-Methode siehe Kapitel 6.2.3.3.

6.2.5.2 Multigruppenanalyse

Zur exakten Überprüfung, ob das Messmodell in der Kalibrierungs- und Validierungsstichprobe gleich bzw. invariant ist, wird eine Multigruppenanalyse durchgeführt (Bollen, 1989; Byrne, 2001; Kline, 1998). Es gibt verschiedene Formen der Invarianz von Messmodellen (Steenkamp & Baumgartner, 1998; Steinmetz, Schmidt, Tina-Booh & Wieczorek, in press; Widaman, 2000): (a) *Konfigurale* Invarianz liegt vor, wenn für alle Gruppen die gleiche Anzahl an Faktoren vorliegt und das Muster von fixierten und freien Faktorladungen gleich ist (Horn & McArdle, 1992); (b) Von *metrischer* Invarianz spricht man, wenn für alle Gruppen die Faktorladungen gleich sind. Bei vorliegender metrischer Invarianz sind die Faktoren inhaltlich zwischen den Gruppen zu vergleichen (Vandenberg & Lance, 2000); (c) *Skalare* Invarianz (Steenkamp & Baumgartner, 1998) liegt vor, wenn die Itemkonstanten (Item-Intercepts) in allen Gruppen gleich sind; (d) Die Invarianz der *Faktor-Kovarianzen* ist gegeben, wenn die Beziehungen zwischen den latenten Faktoren in allen Gruppen gleich sind (Cole & Maxwell, 1985); (e) Die Analyse der Invarianz der latenten Mittelwerte (latent means) impliziert ein Vergleich der latenten Mittelwerte zwischen den Gruppen auf Unterschiede. Der Test auf Unterschiede in den latenten Mittelwerten erfordert, dass sowohl die Faktorladungen (metrische Invarianz) als auch die Itemkonstanten (Skalare Invarianz) zwischen den Gruppen gleich sind (Millsap & Everson, 1991); (f) Eine weniger wichtige Form der Invarianz stellt die Gleichheit der Fehlervarianzen dar (Steinmetz et al., in press).

Mehrere hierarchisch abgestufte Modelle wurden zur Kreuzvalidierung berechnet, wobei ausgehend von einem Modell ohne Restriktion die Modelle immer restriktiver werden:

- *Modell 1.* ohne Restriktion (Alle Parameter können in beiden Gruppen variieren, Basismodell; Test auf konfigurale Invarianz)
- *Modell 2.* Faktorladungen sind in beiden Gruppen gleich (Test auf metrische Invarianz)
- *Modell 3.* wie 2 und Itemkonstanten sind gleich (Test auf skalare Invarianz)
- *Modell 4.* wie 3 und Interkorrelationen sind gleich (Test auf Invarianz der Faktor-Kovarianzen)
- *Modell 5.* wie 4 und Fehlervarianzen sind gleich (Test auf Invarianz der Fehlervarianzen)

Der Test auf Unterschiede in den latenten Mittelwerten erfolgt getrennt (s.u.). Alle Modelle zeigen eine gute Datenanpassung, die Fit-Indizes unterscheiden sich zwischen den Modellen nur marginal (s. Tabelle 14).

Tabelle 14. Fit-Indizes der Modelle der Multigruppenanalyse zur Kreuzvalidierung des Messmodells

Modell	χ^2	df	<i>p</i>	CFI	RMSEA	SRMR	χ^2/df
1	743.23	376	<.001	.94	.04	.06	1.98
2	764.94	391	<.001	.94	.04	.06	1.96
3	784.15	413	<.001	.94	.04	.06	1.90
4	824.76	441	<.001	.94	.04	.06	1.87
5	868.28	463	<.001	.94	.04	.06	1.88

Anmerkungen: *Modell 1.* ohne Restriktion; kein Parameter wurde zwischen beiden Stichproben gleichgesetzt; *Modell 2.* Faktorladungen sind in beiden Gruppen gleich; *Modell 3.* wie 2 und Messintercepts sind gleich; *Modell 4.* wie 3 und Interkorrelationen sind gleich; *Modell 5* wie 4 und Fehlervarianzen sind gleich

Ein genesteter Modellvergleich zeigt, dass sich bei der Annahme, dass Modell 1 gültig ist, Modell 2 keine signifikant schlechtere Datenanpassung aufweist ($\Delta\chi^2 = 21.71$, $df = 15$, $p = .12$). Ebenso zeigt Modell 3 bei Annahme, dass Modell 2 gültig ist, keine schlechtere Datenanpassung ($\Delta\chi^2 = 19.22$, $df = 22$, $p = .63$). Modell 4 ist allerdings signifikant schlechter an die Daten angepasst als Modell 3 ($\Delta\chi^2 = 40.61$, $df = 28$, $p < .06$). Lässt man zu, dass die Korrelationen zwischen der Ereigniseinschätzung und Coping in beiden Stichproben frei variieren kann, so ist Modell 4 nicht mehr signifikant schlechter angepasst als Modell 3 ($\Delta\chi^2 = 34.69$, $df = 27$, $p = .15$). Modell 5 zeigt eine signifikante schlechtere Datenanpassung als das modifizierte Modell 4 ($\Delta\chi^2 = 43.01$, $df = 22$, $p < .01$). Setzt man die Fehlervarianzen des dritten Parcels von Neurotizismus und des dritten Parcels von Selbstwirksamkeitserwartungen frei, so ist die Anpassung des so modifizierten Modells 5 nicht mehr signifikant schlechter als die des Modells 4 ($\Delta\chi^2 = 26.92$, $df = 20$, $p = .14$).

Die Multigruppenanalyse zeigt, dass die inhaltliche Bedeutung der latenten Faktoren zwischen den beiden Stichproben gleich ist (volle Invarianz; vgl. Byrne, Shavelson & Muthén, 1989; Steenkamp & Baumgartner, 1998), da sowohl die Faktorladungen als auch die Ladungskonstanten in beiden Stichproben statistisch nicht voneinander abweichen (Meredith, 1964, 1993). Somit ist es möglich zu prüfen, ob es Unterschiede in den Mittelwerten in den Strukturvariablen zwischen beiden Stichproben gibt. Ein Mittelwertsunterschied würde auf Verzerrungen bei der Zufallsaufteilung hindeuten (vgl. Kap. 6.2.3.1). Hierzu wird auf das Verfahren von Sörbom (1974) zurückgegriffen, welches in AMOS 5.0 automatisch bei Multigruppenanalysen implementiert wird (s. Arbuckle, 2003). Dabei werden in einer Gruppe die

Mittelwerte der Strukturvariablen auf null restringiert (im vorliegenden Fall in der Kalibrierungsstichprobe), in der anderen Gruppe können die Mittelwerte dagegen frei variieren. Wenn ein geschätzter Mittelwert in der Gruppe, in der die Mittelwerte frei geschätzt werden, signifikant von null abweicht, so unterscheiden sich die Mittelwerte zwischen beiden Gruppen signifikant voneinander. Die hier beschriebene Analyse zeigt, dass sich die Mittelwerte in den Strukturvariablen zwischen der Kalibrierungs- und Validierungsstichprobe nicht unterscheiden, in Tabelle 15 sind die Ergebnisse zusammengefasst.

Tabelle 15. Mittelwertsschätzung, Standardmessfehler der Strukturvariablen für die Validierungsstichprobe

Strukturvariable	Schätzer	Messfehler	Signifikanzniveau
Neurotizismus	.01	.13	.90
Soziale Unterstützung	.05	.06	.47
Ereigniseinschätzung	.03	.06	.60
Selbstwirksamkeitserwartungen	.05	.07	.48
Coping	-.05	.06	.53
Beschwerden	-.01	.07	.95
Psychisches Wohlbefinden	-.04	.07	.41

In Tabelle 16 sind die Mittelwerte für beide Stichproben aufgeführt.

Tabelle 16. Mittelwerte in den Strukturvariablen für beide Stichproben

Strukturvariable	Kalibrierungsstichprobe (<i>N</i> = 318)		Validierungsstichprobe (<i>N</i> = 317)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Neurotizismus	3.10	1.03	3.12	1.05
Soziale Unterstützung	3.78	0.77	3.82	0.78
Ereigniseinschätzung	2.33	0.91	2.35	0.93
Selbstwirksamkeitserwartungen	4.30	0.74	4.34	0.73
Coping	4.63	0.72	4.59	0.77
Beschwerden	2.93	0.93	2.93	0.97
Neurotizismus	4.01	0.91	3.97	0.91

Insgesamt lässt sich zusammenfassen, dass das Messmodell sowohl in der Kalibrierungs- als auch in der Validierungsstichprobe eine gute Datenanpassung zeigt. Lediglich in 4 (von 65) Parametern sind signifikante Abweichungen zwischen den beiden Teilstichproben zu bemerken. Dabei kann eine zumindest partielle Kreuzvalidierung des Messmodells konstatiert werden (MacCallum, Roznowski, Mar & Reith, 1994).

6.2.5.3 Zusammenhang der Modellkomponenten mit dem Belastungsausmaß

Da die Modellkomponenten nicht direkt Belastungen wiedergeben, die durch das Referendariat bedingt sind, soll folgend der Zusammenhang zwischen den Modellvariablen und den Belastungsdimensionen analysiert werden. Hierzu werden einfache Korrelationen zwischen den vier Subskalen der Belastungsskala mit den Modellvariablen berechnet. Darüber hinaus wird überprüft, ob soziodemographische Variablen einen Einfluss auf das Belastungsausmaß haben. Im Folgenden wird nur auf die Kalibrierungsstichprobe eingegangen, alle Analysen wurden auch für die Validierungsstichprobe berechnet, die Ergebnisse unterscheiden sich nicht von denen der Kalibrierungsstichprobe.

In Tabelle 17 sind zunächst die Mittelwerte und Standardabweichungen für die Belastungsskalen abgebildet.

Tabelle 17. Mittelwerte und Standardabweichungen für die Belastungsskalen (Kalibrierungsstichprobe)

Belastungsskala	<i>M</i>	<i>SD</i>
Belastungen durch Schülerinnen und Schüler	4.02	1.18
Belastungen durch Konflikte in verschiedenen Bereichen	2.50	0.87
Belastungen durch zukünftige Prüfungen	3.98	1.82
Belastungen durch Unterricht	4.20	0.93

Es zeigt sich, dass am wenigsten Belastungen durch Konflikte empfunden werden, das Belastungsausmaß in den anderen Belastungsquellen ist vergleichbar ausgeprägt, wobei die hohe Standardabweichung für die Belastungsskala „Belastungen durch zukünftige Prüfungen“ auffällig ist. Die zukünftigen Examensprüfungen und die pädagogische Prüfungsarbeit sind für die verschiedenen Ausbildungsphasen in einem unterschiedlichen Ausmaß präsent, die höchsten Belastungen geben ReferendarInnen an, die sich zum Zeitpunkt der Befragung in der Intensivphase befanden.

Die einfachen Korrelationen zwischen den Belastungsskalen und den Modellvariablen fallen, wenn auch von moderater Höhe, zum größten Teil wie erwartet aus (s. Tabelle 18).

Tabelle 18. Einfache Korrelationen zwischen den Belastungsskalen und den Modellvariablen

Belastungen durch...	neuro	sozgr	bewert	wirk	cope	beschw	psych
...Schülerinnen und Schüler	.24***	-.19**	.19**	-.11*	-.04	.08	-.11*
...Konflikte	.30***	-.17**	.34***	-.10	-.19**	.27***	-.31***
...zukünftige Prüfungen	.13*	-.05	.13*	.06	-.05	.15**	-.21***
...durch Unterricht	.44***	-.21***	.46***	-.25***	-.24***	.37***	-.47***

Anmerkungen: neuro = Neurotizismus; sozgr = Soziale Unterstützung; bewert = Ereigniseinschätzung; wirk = Selbstwirksamkeitserwartungen; cope = Coping; beschw = körperliche Beschwerden; psych = psychisches Wohlbefinden

Die engsten Korrelationen weisen die Modellvariablen mit Belastungen durch den Unterricht auf, kaum nennenswerte Korrelationen zeigen sich mit Belastungen durch zukünftige Prüfungen. Ein höheres Ausmaß an (subjektiv) wahrgenommenen Belastungen geht mit mehr

körperlichen Beschwerden und einem geringen psychischen Wohlbefinden einher, wobei Belastungen durch Schülerinnen und Schüler keinen nennenswerten Einfluss auf die Befindlichkeitsindizes haben. Höhere Neurotizismuswerte gehen erwartungsgemäß mit einem höheren Ausmaß an Belastungen in allen vier Bereichen einher, wohingegen bei einem höheren Ausmaß an sozialer Unterstützung das Belastungsausmaß niedriger ausfällt. Je negativer die Ereignisseinschätzung ausfällt, desto mehr Belastungen werden angegeben, die Selbstwirksamkeitserwartungen haben dagegen nur einen geringen Einfluss auf das Belastungsausmaß: Je höher die Selbstwirksamkeitserwartungen, desto geringer das Ausmaß an Belastungen durch den Unterricht oder durch Probleme mit den Schülerinnen und Schülern. Auch das Coping weist nur geringe Zusammenhänge auf, problemzentrierte Strategien gehen mit einem geringeren Belastungsempfinden, ausgelöst durch den Unterricht und Konflikte, einher.

Insgesamt bestätigt das Korrelationsmuster die Annahmen, dass die Variablen des Stressmodells einen Einfluss, wenn auch größtenteils von einem moderaten Ausmaß, auf das subjektive Belastungsausmaß der Befragten haben. Somit stehen die Modellkomponenten im direkten Zusammenhang mit referendariatspezifischen Belastungen.

6.2.5.4 Einfluss von Drittvariablen

Vor der eigentlichen Modellprüfung soll an dieser Stelle untersucht werden, ob Drittvariablen einen Einfluss auf die Modellvariablen haben. Insbesondere soll analysiert werden, ob Geschlecht²⁷, Alter, Familienstand, Kinderlosigkeit (Kinder ja/nein), die Ausbildungsphase und die Studienseminarform (Ausbildung für das Lehramt an Gymnasien, an GHRS oder an beruflichen Schulen) systematisch mit den Modellvariablen zusammenhängen. Bei systematischen Einflüssen muss dies bei der Modellprüfung Berücksichtigung finden.

Zur Überprüfung möglicher Zusammenhänge der Drittvariablen mit den Modellvariablen wurden in einem ersten Schritt einfache Korrelationen für stetige Variablen und einfaktorielle ANOVAs für kategoriale Variablen berechnet, wobei nur auf die Kalibrierungsstichprobe zurückgegriffen wird, da es keine Unterschiede in den aufgezählten Drittvariablen zwischen beiden Stichproben gibt (s.o.).²⁸

²⁷ Gerade für Geschlechtsunterschiede gibt es Hinweise in der Forschungsliteratur. So finden Tamres, Janicki und Helgeson (2002) in einer Meta-Analyse Geschlechtsunterschiede im Bewältigungsverhalten: Frauen nutzen in einem höheren Ausmaß verschiedene Bewältigungsstrategien. Allerdings gibt es auch Evidenz dafür, dass die Geschlechtsunterschiede in Teilen auf ein unterschiedliches Ausmaß in Neurotizismus zurückzuführen sind (Mak, Blewitt & Heaven, 2004).

²⁸ Alle folgenden Analysen wurden zur Kontrolle auch in der Validierungsstichprobe durchgeführt. Es ergaben sich wie erwartet keine Unterschiede in den Ergebnissen.

Insgesamt sind nur wenige signifikante Einflüsse zwischen den Drittvariablen und den Modellvariablen zu konstatieren. Zunächst zeigt sich ein negativer Einfluss des Alters auf das Ausmaß an körperlichen Beschwerden ($r = -.15, p < .01$): Je älter (!) die befragten ReferendarInnen sind, desto weniger körperliche Beschwerden geben sie an. Dieser Befund überrascht insofern, da viele Studien einen umgekehrten Zusammenhang feststellen (z.B. Brähler, Schumacher & Brähler, 2000). Das Alter hat auf alle weiteren Modellvariablen keinen Einfluss und wird aus diesem Grund bei der Modellprüfung nicht weiter berücksichtigt. Weiterhin zeigt sich wie erwartet ein Geschlechtseffekt in der Neurotizismusskala (s. Borkenau & Ostendorf, 1993) und der Beschwerdenskala (s. Nolen-Hoeksema, 2001): Die weiblichen Befragten ($N = 221$) weisen höhere Neurotizismuswerte ($M = 3.21, SD = 1.04$) auf als die männlichen Befragten ($N = 97: M = 2.87, SD = 0.99; F_{1, 316} = 7.22, p < .01$). Auch in der Beschwerdeskala haben die weiblichen Befragten ($M = 3.11, SD = 0.90$) höhere Werte im Vergleich zu den männlichen Befragten ($M = 2.54, SD = 0.88; F_{1, 316} = 26.65, p < .001$).

Auch zwischen den Ausbildungsphasen gibt es bedeutsame Unterschiede, die in Tabelle 19 aufgeführt sind.

Tabelle 19. Mittelwerte und Standardabweichungen für die Ausbildungsphasen der Kalibrierungsstichprobe

	Ausbildungsphase					
	Differenzierungsphase ($N = 101$)		Intensivphase I ($N = 108$)		Intensivphase II ($N = 109$)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Selbstwirksamkeitseinschätzung ($F_{2,315} = 3.10, p < .05$)	4.24 ^a	0.78	4.21 ^a	0.80	4.44 ^b	0.63
Psychisches Wohlbefinden ($F_{2,315} = 3.49, p < .05$)	4.17 ^a	0.84	4.03 ^{a,b}	0.91	3.85 ^b	0.95

Anmerkung: Mittelwerte einer Zeile mit unterschiedlichen Buchstabenindizes unterscheiden sich signifikant voneinander (Scheffé-Test, $p < .05$)

Die Selbstwirksamkeitserwartungen sind in der Intensivphase II höher ausgeprägt als in der Differenzierungsphase und Intensivphase I. Das psychische Wohlbefinden ist dagegen in

der Intensivphase II signifikant niedriger als in der Differenzierungsphase. Keinen Einfluss haben dagegen der Familienstand, Kinderlosigkeit und die Studienseminarform.

Bei der nun folgenden Modellprüfung werden abschließend mögliche moderierende Einflüsse des Geschlechts und der Ausbildungsphase auf die Modellzusammenhänge geprüft.

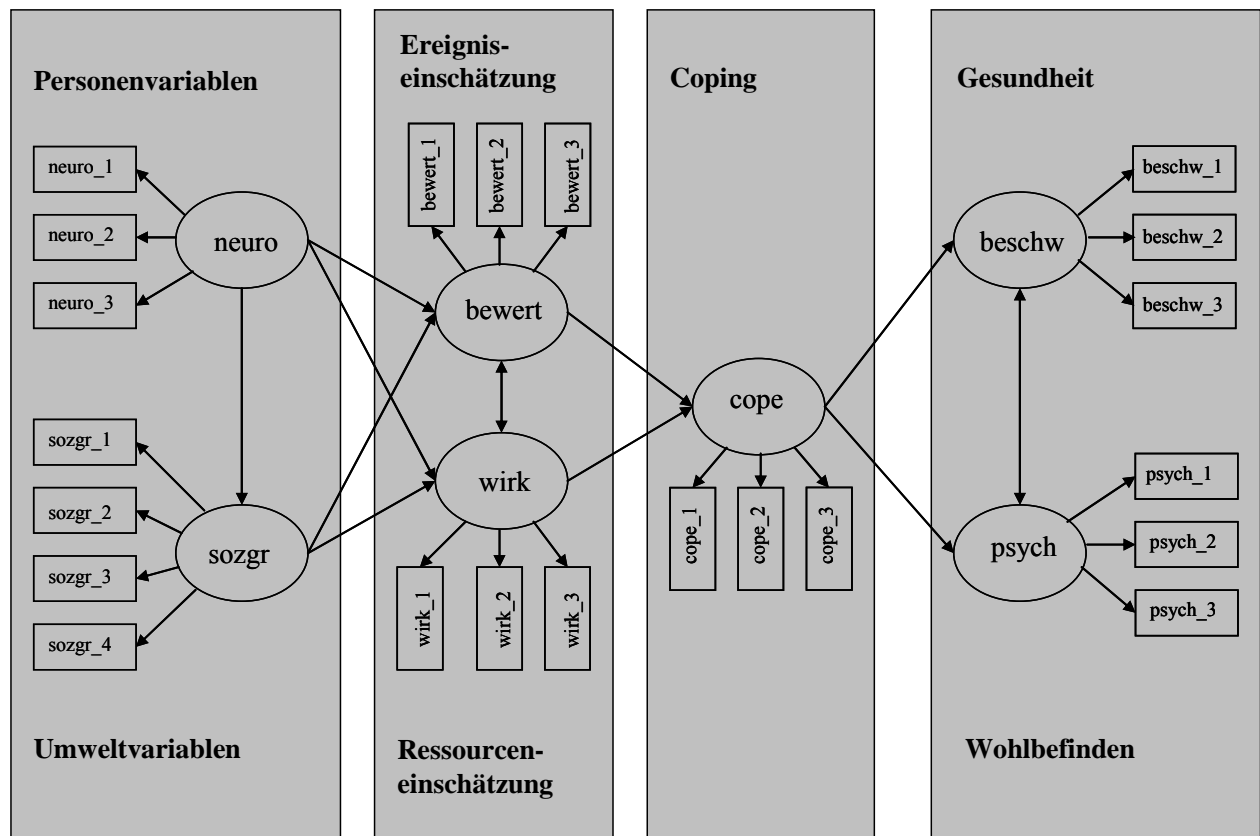
6.2.6 Modellüberprüfung

Im Folgenden wird das im Theorieteil dargestellte Stressmodell überprüft. Dabei wird wiederum in mehreren Schritten vorgegangen. Zunächst wird geprüft, ob das theoretisch postulierte Modell empirisch haltbar ist. Dies ist eine strikte Form der Modellprüfung (MacCallum, 1995; MacCallum & Austin, 2000). In einem weiteren Schritt werden die Daten auf mögliche und notwendige Modifikationen untersucht. Hierbei ist allerdings Vorsicht geboten, da datengetriebene Modellmodifikationen wenig valide (MacCallum, 1986) und zufällig sein können (MacCallum et al., 1992). Aus diesem Grund werden nur inhaltlich plausible Modifikationen vorgenommen, darüber hinaus wird das modifizierte Stressmodell in einem letzten Schritt an der Validierungsstichprobe erneut überprüft, eventuelle Unterschiede sprechen dann für unvalide Modifikationen.

6.2.6.1 Überprüfung des theoretischen Stressmodells

Abbildung 3 enthält das Grundmodell in seiner empirischen Umsetzung, welches mit SEM überprüft werden soll.

Abbildung 3. Das Grundmodell in seiner empirischen Umsetzung



Anmerkungen: neuro = Neurotizismus; sozgr = Soziale Unterstützung; bewert = Ereigniseinschätzung; wirk = Selbstwirksamkeitserwartungen; cope = Coping; beschw = körperliche Beschwerden; psych = psychisches Wohlbefinden

Zunächst ist die Frage zu klären, ob die Ereigniseinschätzung und die Selbstwirksamkeitserwartungen im Sinne der Theorie miteinander interagieren und gemeinsam auf das Bewältigungsverhalten wirken (Dewe, 1991; Jerusalem, 1990; 1991; s. Kap. 3.1).²⁹ Dies wurde vor der eigentlichen Modellüberprüfung unternommen, da die Aufnahme von nicht-linearen Effekten in Strukturgleichungsanalysen - besonders bei komplexen Modellen wie im vorliegenden Fall - problematisch ist (Li et al., 1998). Zur Überprüfung wurde auf eine von Kline und Dunn (2000) vorgeschlagene Methode zurückgegriffen (für Alternativen s. Kenny & Judd, 1984; Ping, 1996a, 1996b; Jaccard & Wan, 1995; Jöreskog & Yang, 1996; für einen Vergleich von Methoden s. Moulder & Algina, 2002 und Yang-Wallentin, Schmidt, Davidov & Bamberg, 2004). Hier ist es in einem ersten Schritt notwendig, die jeweiligen Indikatoren der relevanten Strukturvariablen zu zentrieren (von den jeweiligen individuellen Werten in

²⁹ Mittels zweier Multigruppenanalysen wurde auch geprüft, ob Neurotizismus und das Ausmaß an sozialer Unterstützung mit der Ereigniseinschätzung und Selbstwirksamkeitserwartungen interagieren (vgl. Terry, 1991). Die Analysen zeigen, dass der Einfluss der Ereigniseinschätzung und der Selbstwirksamkeitserwartungen auf das Bewältigungsverhalten in allen Subgruppen (niedrige vs. hohe Neurotizismuswerte; niedriges vs. hohes Ausmaß an sozialer Unterstützung) vergleichbar ist (es ergeben sich in den Multigruppenanalysen keine statistisch bedeutsamen Abweichungen in den Strukturparametern).

den Indikatoren wird der Stichprobenmittelwert abgezogen; vgl. Aiken & West, 1991). Anschließend werden die zentrierten Indikatoren der Strukturvariablen jeweils miteinander multipliziert, die Produkte dienen dann als Indikatoren für den Interaktionsterm. Sowohl für die Kalibrierungs- als auch die Validierungsstichprobe wurde nach diesem Verfahren ein Interaktionsterm gebildet und anschließend jeweils getrennt für beide Stichproben in einer Strukturgleichungsanalyse überprüft, ob es neben den direkten Effekten auch einen Interaktionseffekt gibt. In beiden Fällen, also sowohl in der Kalibrierungs- ($\gamma = -.09, p = .09$) als auch in der Validierungsstichprobe ($\gamma = -.04, p = .42$) ist der Einfluss der Interaktionsstrukturvariable nicht signifikant, nur die beiden direkten Effekte von der Ereigniseinschätzung und Selbstwirksamkeitserwartungen weisen einen statistisch bedeutsamen Effekt auf.³⁰ Somit wird in der weiteren Überprüfung des Modells der Interaktionsterm nicht mit aufgenommen, Hypothese H6 muss verworfen werden.

Das Grundmodell zeigt in der Kalibrierungsstichprobe keine befriedigende Anpassung an die Daten ($\chi^2 = 516.15; df = 198, p < .001, CFI = .90, RMSEA = .07, SRMR = .08, \chi^2/df = 2.61$). Bei der Modellberechnung sind keine Schwierigkeiten aufgetreten, alle Skalenterkorrelationen, Pfade und Fehlervarianzen liegen in einem plausiblen Bereich. Das theoretisch angenommene Modell kann allerdings nicht bestätigt werden, für eine befriedigende Anpassung an die Daten müssen Modifikationen vorgenommen werden.

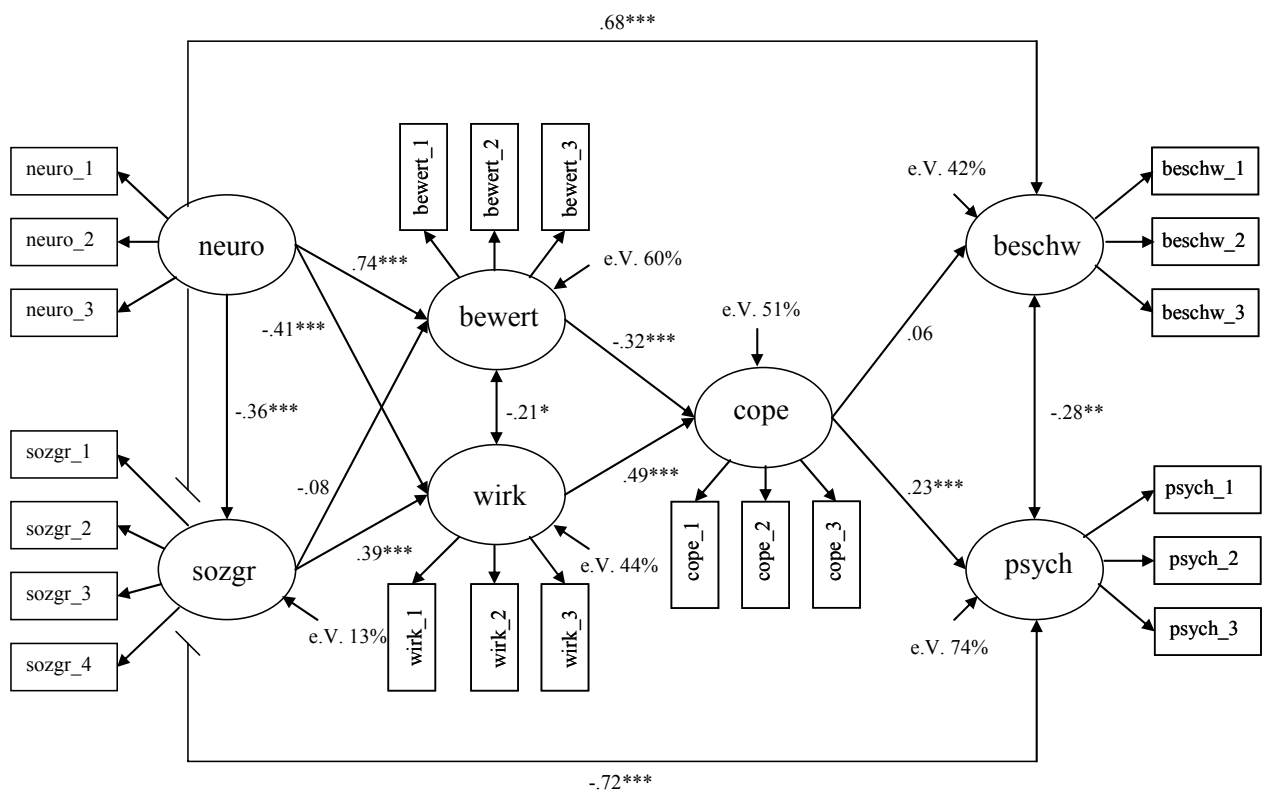
Auf Basis der Modifikationsindizes wird in einem ersten Schritt der Pfad von Neurotizismus auf Psychisches Wohlbefinden zugelassen. Dieses modifizierte Modell ($\chi^2 = 430.31, df = 197, p < .001, CFI = .93, RMSEA = .06, SRMR = .08, \chi^2/df = 2.18$) zeigt eine signifikant bessere Anpassung ($\Delta\chi^2 = 85.84, df = 1, p < .001$) an die Daten als das Grundmodell. Das Modell lässt sich weiter durch die Hinzunahme eines Pfades von Neurotizismus auf körperliche Beschwerden verbessern, der Modell-Fit des so modifizierten Modells ($\chi^2 = 363.12, df = 196, p < .001, CFI = .95, RMSEA = .05, p\text{-close} = .35, SRMR = .05, \chi^2/df = 1.85$) verbessert sich wiederum signifikant ($\Delta\chi^2 = 67.19, df = 1, p < .001$). Nach diesem Modifikationsschritt werden keine weiteren inhaltlich sinnvollen Modifikationen – auf die oben beschriebenen Modifikationen wird weiter unten näher eingegangen - vorgeschlagen. Neurotizismus als relativ stabile Persönlichkeitsvariable hat auch direkte Effekte auf die

³⁰ Wird die 1-Indikator-Methode von Jöreskog und Yang (1996) angewendet, kommt es zu Schwierigkeiten bei der Parameterschätzung (standardisierte ρ -Gewichte > 1). Dies ist mit der Hauptgrund, warum auf die Methode von Kline und Dunn (2000) zurückgegriffen wird. Da gleichzeitig die Probleme mit der anderen Methode auf die Schwierigkeit der Implementierung von nicht-linearen Effekten in Strukturgleichungsmodellen hinweisen, wurde zur Absicherung eine einfache Regressionsanalyse nach der Methode von Aiken und West (1991) berechnet. Auch in dieser Analyse werden nur die direkten Effekte signifikant, der Interaktionsterm in beiden Stichproben erreicht nicht ein konventionelles Signifikanzniveau.

beiden Ergebnisvariablen, und nicht nur indirekte, über die Prozessvariablen vermittelte Einflüsse.

In Abbildung 4 ist das modifizierte Stressmodell abgebildet. Somit lässt sich die im Theorieteil aufgestellte Hypothese 1 (H1) nicht bestätigen. Dieses modifizierte Modell ist nach den gängigen Kriterien (Hu & Bentler, 1999) sehr gut an die Daten angepasst. Alle relevanten Fit-Indizes liegen in einem sehr guten Wertebereich (CFI, RMSEA, SRMR).³¹

Abbildung 4. Das modifizierte Stressmodell mit allen Strukturgleichungen, Korrelationen und der aufgeklärten Varianz der endogenen Faktoren für die Kalibrierungsstichprobe



Anmerkungen: * $p < .05$; ** $p < .01$; *** $p < .001$; e.V. = erklärte Varianz; neuro = Neurotizismus; sozgr = Soziale Unterstützung; bewert = Ereignisseinschätzung; wirk = Selbstwirksamkeitserwartungen; cope = Coping; beschw = körperliche Beschwerden; psych = psychisches Wohlbefinden

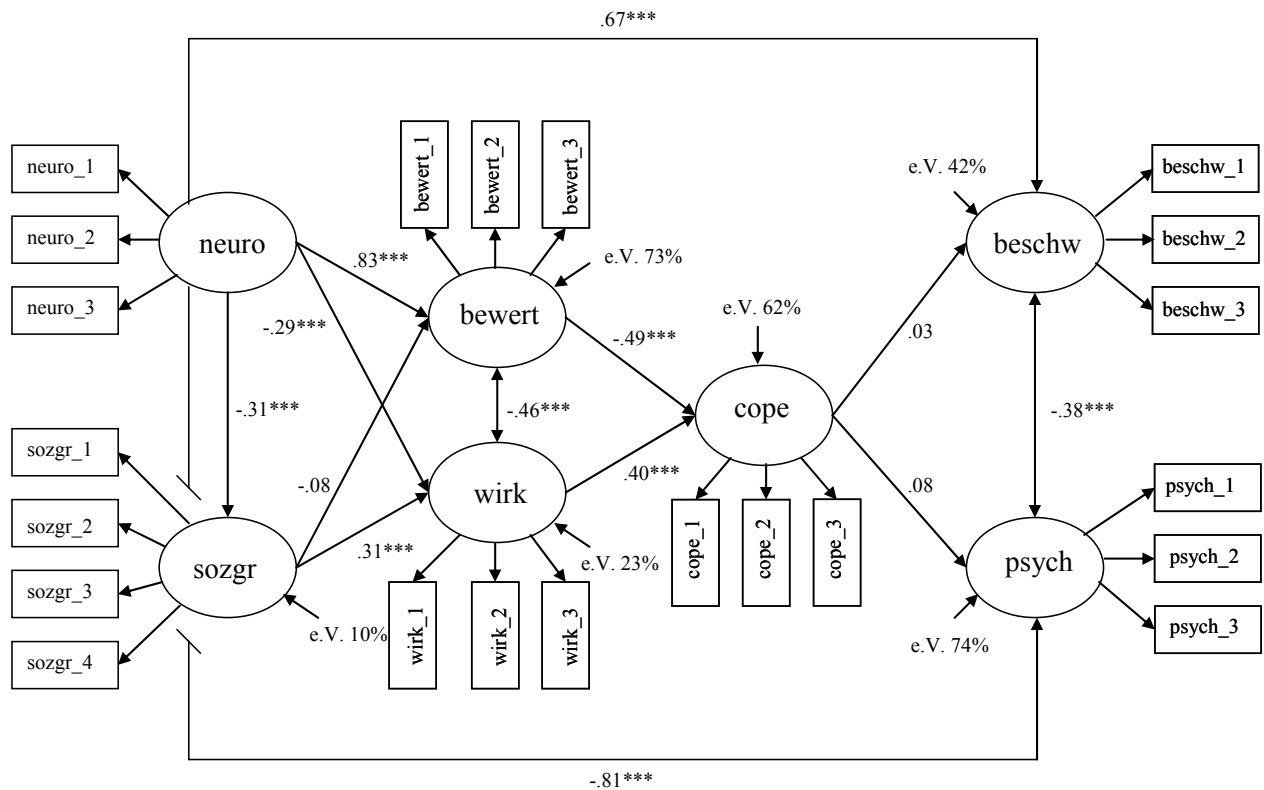
Bei Sichtung der einzelnen strukturellen Zusammenhänge zwischen den latenten Modellvariablen bestätigen sich die Mehrzahl der in Kapitel 5.2 beschriebenen Hypothesen: Höhere Neurotizismuswerte gehen mit einem niedrigeren Ausmaß an sozialer Unterstützung einher

³¹ Eine Power-Analyse (MacCallum et al, 1996; MacCallum & Hong, 1997) ergibt für die Kalibrierungsstichprobe für $df = 196$ und $\alpha = .05$ einen Wert von 1 (gleiches gilt auch für die Validierungsstichprobe). Dieser Wert bedeutet, dass die Wahrscheinlichkeit 1 ist, dass die Hypothese $RMSEA \leq .05$ nicht angenommen wird, wenn der tatsächliche $RMSEA = .08$ ist. Damit ist in der vorliegenden Studie ausreichend statistische Power gegeben.

(H9), mit einer negativeren Ereigniseinschätzung und mit niedrigeren Selbstwirksamkeitserwartungen einher (H3). Umgekehrt führt ein höheres Ausmaß an sozialer Unterstützung zu höheren Selbstwirksamkeitserwartungen (H4). Eine negative Ereigniseinschätzung geht mit der Nutzung emotionszentrierter Bewältigungsstrategien einher, während höhere Selbstwirksamkeitserwartungen mit der Nutzung von mehr problemzentrierten Bewältigungsstrategien korrespondiert (H5). Das Bewältigungsverhalten hat einen Einfluss auf das psychische Wohlbefinden (H8): Problemzentrierte Bewältigungsstrategien bewirken ein besseres psychisches Wohlbefinden. Die Ereigniseinschätzung und die Selbstwirksamkeitserwartungen korrelieren negativ miteinander, ebenso körperliche Beschwerden und psychisches Wohlbefinden (H10). Entgegen den Erwartungen beeinflusst das Ausmaß an sozialer Unterstützung nicht die Ereigniseinschätzung (H5), die Nutzung von mehr problemzentrierten Bewältigungsstrategien reduziert nicht das Ausmaß an körperlichen Beschwerden (H7). Die Mehrzahl der Hypothesen lassen sich aber in Studie 1 bestätigen.

Auch in der Validierungsstichprobe zeigt das modifizierte Stressmodell eine befriedigende Datenanpassung: ($\chi^2 = 453.24$, $df = 196$, $p < .001$, $CFI = .93$, $RMSEA = .06$, $p\text{-close} < .01$, $SRMR = .06$, $\chi^2/df = 2.31$). Ein erster augenscheinlicher Parametervergleich zeigt, dass es nur wenige auffallende Abweichungen zwischen den beiden Stichproben in den Parameterschätzern gibt: Hauptunterschied ist, dass in der Validierungsstichprobe Coping keinen signifikanten Zusammenhang mit psychischem Wohlbefinden hat (vgl. Abbildung 5).

Abbildung 5. Das modifizierte Stressmodell mit allen Strukturgleichungen, Korrelationen und der aufgeklärten Varianz der endogenen Faktoren für die Validierungsstichprobe



Anmerkungen: * $p < .05$; ** $p < .01$; *** $p < .001$; e.V. = erklärte Varianz; neuro = Neurotizismus; sozgr = Soziale Unterstützung; bewert = Ereignisseinschätzung; wirk = Selbstwirksamkeitserwartungen; cope = Coping; beschw = körperliche Beschwerden; psych = psychisches Wohlbefinden

Die Stabilität der Parameter für das modifizierte Stressmodell wurde wie für das Messmodell für beide Stichproben mit der Bootstrap-Methode überprüft. Für beide Stichproben zeigt sich, dass die prozentuale Unterschätzung der Standardfehler für die Modellparameter in einem akzeptablen Bereich liegt. Sowohl die prozentuale Abweichung der Standardfehler für die Faktorladungen (Kalibrierungsstichprobe: $M = 13.96\%$, $SD = 16.57$, $Min = -6.38$, $Max = 53.88$; Validierungsstichprobe: $M = 9.34\%$, $SD = 12.90$, $Min = -12.68$, $Max = 35.66$) als auch für die Strukturgleichungen und Skaleninterkorrelationen (Kalibrierungsstichprobe: $M = 25.78\%$, $SD = 30.57$, $Min = -4.17$, $Max = 109.46$; Validierungsstichprobe: $M = 21.48\%$, $SD = 11.33$, $Min = 0.00$, $Max = 40.00$) sind kaum erhöht, der Bias hält sich sowohl für die Faktorladungen (Kalibrierungsstichprobe: $M = .009$, $SD = .006$, $Min = .003$, $Max = .020$; Validierungsstichprobe: $M = .007$, $SD = .003$, $Min = .003$, $Max = .012$) als auch für die Strukturgleichungen und Skaleninterkorrelationen (Kalibrierungsstichprobe: $M = .002$, $SD = .008$, $Min = -6.38$, $Max = 53.88$; Validierungsstichprobe: $M = 9.34$, $SD = 12.90$, $Min = -12.68$,

Max = 35.66) in Grenzen. Auch bei den bias-korrigierten Konfidenzintervallen für die Strukturgewichte und Interkorrelationen ist bei Parametern, die mit der ML-Methode statistisch signifikant sind, die Null nicht in den Intervallen eingeschlossen. Allerdings kann in der Kalibrierungsstichprobe für den Pfad von sozialer Unterstützung auf die Ereignisseinschätzung nicht ausgeschlossen werden, dass in der Population kein Zusammenhang besteht. Gleiches gilt für die Korrelation zwischen der Ereignisseinschätzung und den Selbstwirksamkeitserwartungen. In beiden Fällen umschließt das Konfidenzintervall die Null. In der Validierungsstichprobe dagegen bestätigt die Bootstrap-Methode die Ergebnisse der Analysen mit der ML-Methode. Insgesamt kann auch hier wie für das Messmodell resümiert werden, dass die mögliche Abweichung der Daten von der multivariaten Normalverteilung zu keiner massiven Unterschätzung der Standardfehler führt und die statistische Bedeutsamkeit von Parametern nicht überschätzt wird.

Die Modellberechnung basierend auf der FIML-Methode (s.o.) kommt sowohl in der Kalibrierungs- wie auch Validierungsstichprobe zu gleichen Ergebnissen. Wie schon für das Messmodell kann auch für die Modellüberprüfung eine Verzerrung der Ergebnisse verursacht durch unterschiedliche Methoden zum Umgang mit fehlenden Werten weitgehend ausgeschlossen werden.

Da von keinen Verzerrungen in den Parameterschätzern ausgegangen werden kann, liefert auch die folgende Kreuzvalidierung keine verzerrten Ergebnisse. Unterschiede, die festgestellt werden, können nicht auf fehlerhafte Parameterschätzer zurückgeführt werden.

6.2.6.2 Kreuzvalidierung

Zur Kreuzvalidierung des empirisch modifizierten Stressmodells wird auf das gleiche Verfahren zurückgegriffen wie bei der Kreuzvalidierung des Messmodells (s.o.). D.h. es werden mehrere hierarchische, genestete Modelle berechnet. Basierend auf einem Modell, in dem kein Parameter zwischen beiden Stichproben gleichgesetzt wurde (Modell ohne Restriktion) werden sukzessive die Modelle restriktiver:

- *Modell 1.* ohne Restriktion; kein Parameter wurde zwischen beiden Stichproben gleichgesetzt
- *Modell 2.* Faktorladungen sind in beiden Gruppen gleich
- *Modell 3.* wie 2 und Strukturgewichte sind gleich
- *Modell 4.* wie 3 und Interkorrelationen sind gleich
- *Modell 5.* wie 4 und Residualvarianz der Strukturvariablen ist gleich

- *Modell 6.* wie 5 und Fehlervarianzen sind gleich bis auf das dritte Parcel von Neurotizismus und Selbstwirksamkeitserwartungen

Das Vorgehen orientiert sich dabei an den Empfehlungen von Bollen (1989), Byrne (2001) und Kline (1998). Alle Modelle zeigen eine gute Datenanpassung, die Fit-Indizes unterscheiden sich zwischen den Modellen nur marginal (s. Tabelle 20).

Tabelle 20. Fit-Indizes der Modelle der Multigruppenanalyse

Modell	χ^2	df	<i>p</i>	CFI	RMSEA	SRMR	χ^2/df
1	816.37	392	<.001	.94	.04	.05	2.08
2	833.79	407	<.001	.94	.04	.06	2.05
3	847.34	418	<.001	.94	.04	.06	2.03
4	847.60	419	<.001	.94	.04	.06	2.02
5	860.26	427	<.001	.94	.04	.06	2.02
6	888.60	447	<.001	.93	.04	.06	1.99

Anmerkungen: *Modell 1.* ohne Restriktion; kein Parameter wurde zwischen beiden Stichproben gleichgesetzt; *Modell 2.* Faktorladungen sind in beiden Gruppen gleich; *Modell 3.* wie 2 und Strukturgewichte sind gleich; *Modell 4:* wie 3 und Interkorrelationen sind gleich; *Modell 5.* wie 4 und Residualvarianz der Strukturvariablen ist gleich und Interkorrelationen sind gleich; *Modell 6* wie 5 und Fehlervarianzen sind gleich

Ein genesteter Modellvergleich zeigt, dass sich bei Annahme, dass Modell 1 gültig ist, Modell 2 keine signifikant schlechtere Datenanpassung aufweist ($\Delta\chi^2 = 17.43$, $df = 15$, $p = .29$). Ebenso zeigt Modell 3 bei Annahme, dass Modell 2 gültig ist, keine schlechtere Datenanpassung ($\Delta\chi^2 = 13.55$, $df = 11$, $p = .26$). Auch Modell 4 ($\Delta\chi^2 = 0.26$, $df = 1$, $p = .61$), Modell 5 ($\Delta\chi^2 = 12.66$, $df = 8$, $p = .12$) und Modell 6 ($\Delta\chi^2 = 28.34$, $df = 20$, $p = .10$) sind jeweils nicht signifikant schlechter an die Daten angepasst unter der Annahme, dass die jeweiligen Vorgängermodelle gültig sind. Zusammenfassend zeigt die Multigruppenanalyse, dass das in der Kalibrierungsstichprobe empirisch modifizierte Stressmodell erfolgreich an der Validierungsstichprobe kreuzvalidiert werden kann. Es gibt keine statistisch bedeutsamen Abweichungen in den Modellparametern.

Allerdings zeigen sich bei näherer Betrachtung der Parameterschätzer bedeutsame Unterschiede: So wird, wie bereits oben erwähnt, der Pfad von Coping auf psychisches Wohlbefinden in der Validierungsstichprobe nicht signifikant. Lässt man diesen Pfad frei variieren in der Multigruppenanalyse, so ist die Datenanpassung von Modell 3 signifikant besser als zuvor

($\Delta\chi^2 = 4.06$, $df = 1$, $p < .05$), auch der Unterschied zu Modell 2 wird geringer ($\Delta\chi^2 = 9.49$, $df = 10$, $p = .49$).

6.2.6.3 Äquivalente Modelle

Eine bekannte Problematik bei SEM stellen mathematisch äquivalente Modelle dar (Lee & Hershberger, 1990; MacCallum, Wegener, Uchino & Fabrigar, 1993; Stelzl, 1986). Für jedes gegebene Modell existieren äquivalente Modelle, die ein anderes Muster an Beziehungen zwischen den Variablen modellieren, aber dieselbe Passung auf die Stichprobendaten aufweisen. Somit kann nicht auf Basis statistischer Kennwerte zwischen solchen Modellen unterschieden werden. Auch für das vorliegende Modell gibt es eine Vielzahl von äquivalenten Modellen.³² Beispielsweise kann der Pfad zwischen Neurotizismus und sozialer Unterstützung in seiner Richtung umgekehrt werden, der resultierende Modell-Fit ist derselbe wie im Ursprungsmodell. Kritisch zu diskutieren ist an dieser Stelle allerdings nur ein Modell³³: Statt einer korrelativen Beziehung zwischen den Selbstwirksamkeitserwartungen und der Bedrohungsbewertung kann ein Modell mit einem gerichteten Pfad von den Selbstwirksamkeitserwartungen auf die Ereigniseinschätzung als konkurrierendes Modell nicht ausgeschlossen werden. Es ist durchaus möglich, dass Selbstwirksamkeitserwartungen im Sinne einer personalen Ressource (Schwarzer, 2000) einer negativen Ereigniseinschätzung im Sinne von einer Bedrohungs- und/oder Verlustbewertung vorbeugen. Dagegen spricht die gewählte Operationalisierung der Selbstwirksamkeitserwartungen in ihrer situationsspezifischen Ausformung. Da die Daten der vorliegenden Untersuchung allerdings querschnittlich erhoben sind und die Ereigniseinschätzung nicht situationsspezifisch erhoben wurde, scheint es sinnvoll, lediglich eine korrelative Beziehung zwischen beiden Modellvariablen anzunehmen. Ein solches konkurrierendes Modell kann nur durch experimentelle oder Längsschnittuntersuchungen ausgeschlossen werden (MacCallum et al., 1993).

6.2.6.4 Moderierende Drittvariablen

Bei der Überprüfung des Messmodells wurde gezeigt, dass das Geschlecht wie auch die Ausbildungsphase einen Einfluss auf die Modellvariablen hat. Um mögliche moderierende

³² Stelzl (1986) und Lee und Hershberger (1990) stellen einfache Konstruktionsregeln vor, mit deren Hilfe äquivalente Modelle konstruiert werden können.

³³ Alle weiteren äquivalenten Modelle beinhalten Pfade, die einerseits theoretisch keine Logik aufweisen und andererseits empirisch widerlegt sind.

Wirkungen dieser Variablen zu prüfen, wird das modifizierte Stressmodell getrennt für die weiblichen und männlichen Befragten sowie für die drei Ausbildungsphasen mit SEM berechnet und anschließend wird mit Multigruppenanalysen geprüft, ob die Modelle sich in den Modellparametern unterscheiden. Auch diese Analysen werden nur für die Kalibrierungsstichprobe präsentiert.³⁴

Die Analysen zeigen, dass das modifizierte Stressmodell auch in den Substichproben befriedigende Fit-Werte aufweist, die Ergebnisse für die einzelnen Stichproben sind in Tabelle 21 zusammengefasst.

Tabelle 21. Fit-Werte des modifizierten Stressmodells für die weiblichen und männlichen Befragten sowie getrennt für die Ausbildungsphasen (Kalibrierungsstichprobe)

Substichprobe	Fit-Werte
Frauen ($N = 221$)	$\chi^2 = 306.57$, $df = 196$, $p < .001$, $CFI = .95$, RMSEA = .05, SRMR = .06, $\chi^2/df = 1.56$
Männer ($N = 97$)	$\chi^2 = 268.54$, $df = 196$, $p < .001$, $CFI = .93$, RMSEA = .06, SRMR = .07, $\chi^2/df = 1.37$
Differenzierungsphase ($N = 101$)	$\chi^2 = 278.62$, $df = 196$, $p < .001$, $CFI = .92$, RMSEA = .07, SRMR = .08, $\chi^2/df = 1.42$
Intensivphase I ($N = 108$)	$\chi^2 = 278.05$, $df = 196$, $p < .001$, $CFI = .94$, RMSEA = .06, SRMR = .07, $\chi^2/df = 1.42$,
Intensivphase II ($N = 109$)	$\chi^2 = 263.39$, $df = 196$, $p < .001$, $CFI = .93$, RMSEA = .06, SRMR = .07, $\chi^2/df = 1.34$,

Die Multigruppenanalysen für den Vergleich der weiblichen und männlichen Befragten und die drei Ausbildungsphasen erfolgen nach dem gleichen Prinzip. Es werden mehrere hierarchische, genestete Modelle berechnet, die in ihren Annahmen sukzessive restriktiver werden:

- *Modell 1.* ohne Restriktion; kein Parameter wurde zwischen beiden Stichproben gleichgesetzt

³⁴ Die Analysen wurden auch für die Validierungsstichprobe durchgeführt und kommen zu den gleichen Ergebnissen.

- *Modell 2.* Faktorladungen sind in beiden Gruppen gleich
- *Modell 3.* wie 2 und Strukturgewichte sind gleich
- *Modell 4.* wie 3 und Residualvarianz der Strukturvariablen ist gleich und Interkorrelationen sind gleich
- *Modell 5.* wie 4 und Fehlervarianzen sind gleich

Die Multigruppenanalyse zum Vergleich der weiblichen und männlichen Befragten zeigen zunächst, dass alle Modelle eine gute Datenanpassung aufweisen (s. Tabelle 22).

Tabelle 22. Fit-Werte der Modelle der Multigruppenanalyse zum Vergleich der weiblichen und männlichen Befragten der Kalibrierungsstichprobe

Modell	Fit-Werte
1	$\chi^2 = 575.66$, $df = 393$, $p < .001$, CFI = .94, RMSEA = .04, SRMR = .06, $\chi^2/df = 1.47$
2	$\chi^2 = 597.52$, $df = 407$, $p < .001$, CFI = .94, RMSEA = .04, SRMR = .06, $\chi^2/df = 1.47$
3	$\chi^2 = 608.77$, $df = 418$, $p < .001$, CFI = .94, RMSEA = .04, SRMR = .06, $\chi^2/df = 1.46$
4	$\chi^2 = 618.68$, $df = 427$, $p < .001$, CFI = .94, RMSEA = .06, SRMR = .06, $\chi^2/df = 1.45$
5	$\chi^2 = 646.17$, $df = 449$, $p < .001$, CFI = .94, RMSEA = .06, SRMR = .06, $\chi^2/df = 1.44$

Anmerkungen: *Modell 1.* ohne Restriktion; kein Parameter wurde zwischen beiden Stichproben gleichgesetzt; *Modell 2.* Faktorladungen sind in beiden Gruppen gleich; *Modell 3.* wie 2 und Strukturgewichte sind gleich; *Modell 4.* wie 3 und Residualvarianz der Strukturvariablen ist gleich und Interkorrelationen sind gleich; *Modell 5* wie 4 und Fehlervarianzen sind gleich

Der sukzessive Vergleich der Modelle zeigt darüber hinaus, dass es keine statistisch bedeutsamen Unterschiede in den Modellparametern zwischen den weiblichen und männlichen Befragten gibt, so dass das Geschlecht keine moderierende Wirkung auf die Zusammenhänge zwischen den Modellvariablen hat.

Auch die Modelle der Multigruppenanalyse zum Vergleich der Ausbildungsphasen weisen befriedigende Fit-Werte auf (s. Tabelle 23).

Table 23. Fit-Werte der Modelle der Multigruppenanalyse zum Vergleich der Ausbildungsphasen der Kalibrierungsstichprobe

Modell	Fit-Werte
1	$\chi^2 = 820.01$, $df = 588$, $p < .001$, CFI = .93, RMSEA = .04, SRMR = .08, $\chi^2/df = 1.40$
2	$\chi^2 = 854.49$, $df = 618$, $p < .001$, CFI = .93, RMSEA = .04, SRMR = .08, $\chi^2/df = 1.38$
3	$\chi^2 = 902.53$, $df = 640$, $p < .001$, CFI = .92, RMSEA = .04, SRMR = .09, $\chi^2/df = 1.41$
4	$\chi^2 = 920.71$, $df = 658$, $p < .001$, CFI = .94, RMSEA = .06, SRMR = .09, $\chi^2/df = 1.40$
5	$\chi^2 = 970.16$, $df = 702$, $p < .001$, CFI = .94, RMSEA = .06, SRMR = .09, $\chi^2/df = 1.38$

Anmerkungen: *Modell 1.* ohne Restriktion; kein Parameter wurde zwischen beiden Stichproben gleichgesetzt; *Modell 2.* Faktorladungen sind in beiden Gruppen gleich; *Modell 3.* wie 2 und Strukturgewichte sind gleich; *Modell 4.* wie 3 und Residualvarianz der Strukturvariablen ist gleich und Interkorrelationen sind gleich; *Modell 5* wie 4 und Fehlervarianzen sind gleich

Allerdings zeigt sich, dass Modell 3 signifikant schlechter an die Daten angepasst ist als Modell 2 ($\Delta\chi^2 = 48.04$, $df = 22$, $p < .01$) und es somit Unterschiede in den Strukturgewichten gibt. Die Ausbildungsphasen unterscheiden sich in den Strukturgewichten zweier Pfade: Dem Pfad von Neurotizismus auf die Bedrohungsbewertung (Differenzierungsphase: $\gamma = -.58$, $p < .001$; Intensivphase I: $\gamma = -.65$, $p < .001$; Intensivphase II: $\gamma = -.07$, $p = .59$) und von der Ressourcenbewertung auf Coping (Differenzierungsphase: $\beta = .36$, $p < .05$; Intensivphase I: $\beta = 1.24$, $p < .01$; Intensivphase II: $\beta = .25$, $p = .09$). Lässt man beide Parameter in den Substichproben frei variieren, so ist Modell 3 nicht länger schlechter angepasst als Modell 2 ($\Delta\chi^2 = 26.30$, $df = 18$, $p = .09$). In der Intensivphase II hat Neurotizismus im Gegensatz zu den beiden anderen Ausbildungsphasen keinen Einfluss auf die Bedrohungsbewertung, in der Differenzierungsphase weist die Ressourceneinschätzung einen größeren Effekt auf Coping auf als in den beiden anderen Phasen. Die Ergebnisse zeigen, dass es zwischen den Ausbildungsphasen nur geringe Unterschiede in der Beziehung zwischen den Modellkomponenten gibt.

6.2.6.5 Zusammenfassendes Fazit der Ergebnisse aus Studie 1

Folgend sollen die Ergebnisse in Hinblick auf die in Kapitel 5.2 formulierten Hypothesen diskutiert werden. Hypothese H2 kann in Studie 1 nicht überprüft werden, da das Querschnittsdesign eine Analyse von intraindividuellen Veränderungen und rekursiven Beziehungen über die Zeit nicht ermöglicht.

Das Messmodell zeigt, dass sich nicht alle Konstrukte hinsichtlich ihrer theoretischen Struktur empirisch entsprechend wieder finden lassen, was wiederum Rückwirkungen auf einen Teil der spezifischen Hypothesen hat. So lässt sich die Ereignisseinschätzung in dieser Stichprobe nur mit einem bipolaren Faktor mit den Endpolen Herausforderung und Verlust/Bedrohung abbilden. Auch der Copingskala unterliegt nur ein Faktor, ebenfalls bipolar mit den Endpolen emotionszentrierte vs. problemzentrierte Bewältigung.

Die generelle Struktur des Stressmodells kann nur mit Abstrichen bestätigt werden (H1): Zwar haben Neurotizismus und soziale Unterstützung als Personenmerkmale einen Einfluss auf die Ereignis- und Ressourceneinschätzung (H3, H4). Die Ereignis- und Ressourceneinschätzung wiederum beeinflussen das Bewältigungsverhalten (H5). Das Bewältigungsverhalten hat nur einen Effekt auf das psychische Wohlbefinden (H8), das Ausmaß an körperlichen Beschwerden kann nicht durch das Bewältigungsverhalten erklärt werden (H7). Auch der direkte Pfad von Neurotizismus auf das psychische Wohlbefinden und das Ausmaß an körperlichen Beschwerden wurde nicht erwartet.

Gemäß H9 wird der Pfad von Neurotizismus auf das Ausmaß an sozialer Unterstützung signifikant. Das Ausmaß an körperlichen Beschwerden und das psychische Wohlbefinden korrelieren wie erwartet (H10) miteinander: Ein geringeres Ausmaß an körperlichen Beschwerden geht mit einem besseren psychischen Wohlbefinden einher.

6.3 *Diskussion Studie 1*

Die Ergebnisse von Studie 1 zeigen, dass sich die Struktur der transaktionalen Stresstheorie von Lazarus (Lazarus, 1991a; Lazarus & Folkman, 1984) empirisch nur in Teilen bestätigen lässt. Zwar haben personale und situationale Ressourcen einen Einfluss auf die Ereignis- und Ressourceneinschätzung, die wiederum gemeinsam das Bewältigungsverhalten bestimmen. Das Bewältigungsverhalten hat aber lediglich auf das psychische Wohlbefinden einen Einfluss, körperliche Beschwerden sind in dieser Studie unabhängig von der Art und Weise, wie eine Person mit einer stressigen Situation umgeht. Allerdings zeigen sich auch in früheren Studien ähnliche Befunde, in der Untersuchung von Folkman, Lazarus, Gruen und DeLongis (1986) hat das Bewältigungsverhalten ebenfalls keinen Einfluss auf körperliche Symptome. Neurotizismus hat entgegen den Erwartungen auch direkte Effekte auf die Ergebnisvariablen.

Auch die spezifischen Hypothesen lassen sich nur zum Teil bestätigen. Das Ausmaß an sozialer Unterstützung hat entgegen den Erwartungen keinen Einfluss auf die Ereignisseinschätzung. Dies kann allerdings daran liegen, dass Neurotizismus einen starken Einfluss auf diese Modellkomponente hat und auch negativ mit dem Ausmaß an sozialer Unterstützung

zusammenhängt, was ein Hinweis auf die verzerrte Wahrnehmung von Personen mit hohen Neurotizismuswerten sein kann (vgl. Bolger & Eckenrode, 1991; Kahn & Hessling, 2001). Es bleibt abzuwarten, ob in Studie 2 ähnliche Effekte auftreten. Eine wichtige Abweichung von den Vorhersagen der transaktionalen Stresstheorie sind auch die direkten Effekte von Neurotizismus auf das psychische Wohlbefinden und das Ausmaß an körperlichen Beschwerden. D.h. die Kernvariablen der transaktionalen Stresstheorie, die kognitive Bewertung und das Bewältigungsverhalten, können nur zum Teil die Effekte des stabilen Persönlichkeitsmerkmals Neurotizismus erklären, kognitive Prozesse spielen hier nur eine untergeordnete Rolle. Bolger und Zuckerman (1995) zeigen anhand verschiedener Modelle auf, warum Persönlichkeitsmerkmale einen Einfluss auf das Wohlbefinden und auf Beschwerden haben können. So ist eine Möglichkeit, dass Personen mit hohen Neurotizismuswerten maladaptive Bewältigungsstrategien wählen oder aber generell das Bewältigungsverhalten ineffektiver ist als bei Personen mit niedrigen Neurotizismuswerten. Die erste Möglichkeit trifft in dem getesteten Modell indirekt zu (vermittelt über die kognitive Bewertung nutzen Personen mit höheren Neurotizismuswerten eher emotionszentrierte Bewältigungsstrategien), die zweite Möglichkeit wurde ebenfalls getestet (Moderatoreffekt von Neurotizismus, Vergleich von Personen mit niedrigen und hohen Neurotizismuswerten). So kann an dieser Stelle nicht endgültig geklärt werden, warum starke direkte Effekte von Neurotizismus bleiben. Eine Möglichkeit besteht allerdings darin, dass Personen mit höheren Neurotizismuswerten generell eher über ein schlechtes Wohlbefinden und mehr körperliche Beschwerden klagen, während dieser Bias bei den übrigen Maßen nicht so stark durchschlägt. Eine mögliche Erklärung für ein solches Muster liefern Befunde von Larsen (1992). Er kann zeigen, dass der hohe Zusammenhang zwischen Neurotizismus und psychischen und somatischen Symptomen zum Teil auf Verzerrungen in der Enkodier- und Erinnerungsphase rückführbar ist und hierdurch Personen mit hohen Neurotizismuswerten, unabhängig vom Umgang mit stressigen Situationen, ihre Beschwerden als schwerer erinnern als Personen mit niedrigen Neurotizismuswerten.

Schließlich müssen die Ergebnisse insgesamt als vorläufig gelten aufgrund der im Theorierteil genannten Probleme von Querschnittsstudien. Nur strukturelle Annahmen der transaktionalen Stresstheorie können mit Studie 1 überprüft werden. Das Querschnittsdesign erlaubt nicht die Analyse der prozessualen Annahmen. Studie 2 der vorliegenden Arbeit soll dieses Manko aufheben.

7 Studie 2

Wie bei Darstellung von Studie 1 wird auch bei Studie 2 zunächst die Methode beschrieben (Kap. 7.1), im Anschluss daran werden die Ergebnisse dargestellt (Kap. 7.2) und abschließend diskutiert (Kap. 7.3).

7.1 Methode Studie 2

7.1.1 Vorgehen

Zur Durchführung von Studie 2 wurde auf dieselben Studienseminare zurückgegriffen wie in Studie 1, mit Ausnahme eines Studienseminars für das Lehramt an beruflichen Schulen³⁵, die Stichproben überschneiden sich aber nicht. Für Studie 2 wurde nur auf die ReferendarInnen zurückgegriffen, die zum ersten Messzeitpunkt mit dem Referendariat begonnen hatten (Ausbildungsbeginn November 2001). An Studie 1 beteiligten sich ReferendarInnen, die sich in einem späteren Ausbildungsabschnitt befanden (Ausbildungsbeginn Mai 2001 und früher). Hier konnten aus organisatorischen Gründen keine Daten für die Längsschnittuntersuchung erhoben werden, da dieses Studienseminar erst nach Beginn von Studie 2 kontaktiert wurde und zu diesem Zeitpunkt bereits zwei Befragungen durchgeführt worden waren. Somit beteiligten sich an Studie 2 insgesamt 16 Studienseminare, wovon acht Studienseminare ReferendarInnen für das Lehramt an Grund-, Haupt, Real- und Sonderschulen, sechs für das Lehramt an Gymnasien und zwei für das Lehramt an beruflichen Schulen ausbilden.³⁶ Über einen Zeitraum von insgesamt 13 Monaten wurden insgesamt vier Befragungen durchgeführt³⁷, in Tabelle 24 sind der zeitliche Ablauf und der zu den jeweiligen Messzeitpunkten gegebene Ausbildungsstand der Befragten aufgeführt.

³⁵ Es handelt sich dabei um das Studienseminar, welches zur Nacherhebung von Daten zu einem späteren Zeitpunkt kontaktiert wurde.

³⁶ Im Laufe der Untersuchungen sind je zwei Studienseminare zu einem Studienseminar zusammengelegt worden, die Untersuchungsdurchführung wurde hierdurch nicht beeinträchtigt.

³⁷ Es wurde ein weiterer Messzeitpunkt realisiert, der allerdings nicht in den Rahmen der vorliegenden Studie eingebettet war. Grundlage dieser Befragung bildete eine stark gekürzte Fragebogenversion, Hauptziel war es, die Bereitschaft und Kontaktierungsmöglichkeiten der Befragten zu ermitteln, an späteren Befragungen teilzunehmen, die den Berufseintritt als Untersuchungsgegenstand haben sollen.

Tabelle 24. Zeitlicher Ablauf von Studie 2 und Ausbildungsstand der befragten ReferendarInnen zu den jeweiligen Messzeitpunkten

Messzeitpunkt (Zeitraum)	Ausbildungsstand
t1 (November / Dezember 2001)	Einführungsphase
t2 (Februar 2002)	Beginn der Differenzierungsphase
t3 (Juni 2002)	Ende Differenzierungsphase
t4 (Dezember 2002)	Mitte Intensivphase I

Die Messzeitpunkte wurden so gelegt, dass sie kritische Abschnitte in der Ausbildung bzw. Übergänge zwischen verschiedenen Ausbildungsphasen abdecken. Der erste Messzeitpunkt (t1) diente zur Ermittlung der Baseline, d.h. hier sollte primär erfasst werden, mit welchen persönlichen Ressourcen, körperlichen Beschwerden und psychischem Wohlbefinden die Befragten in das Referendariat eintraten. Messzeitpunkt 2 (t2) wurde auf den Beginn der Differenzierungsphase gelegt, da hier die ReferendarInnen zum ersten Mal eigenständig Unterrichtsstunden leiten mussten. Messzeitpunkt 3 (t3) lag am Ende der Differenzierungsphase und im vorliegenden Fall etwa drei bis vier Wochen vor den Sommerferien (Beginn der Sommerferien in Hessen im Jahr 2002 war der 27. Juni). Zu diesem Zeitpunkt häufen sich erfahrungsgemäß die letzten Klausuren und die Befragten mussten zum ersten Mal Noten vergeben. Der vierte Messzeitpunkt (t4) lag Anfang Dezember und somit in der Mitte der Intensivphase I. Die Intensivphase I geht mit einer Erhöhung des Lehrdeputats einher, somit ist mit einer weiteren Belastungserhöhung zu rechnen (s.a. Kap. 5.1.1).

Zu t1 konnten an allen Studienseminaren die Fragebogen direkt an die ReferendarInnen verteilt werden, zum Großteil wurden die Fragebogen direkt vor Ort ausgefüllt. Dies war möglich, da die ReferendarInnen zu Beginn ihrer Ausbildung fast ausschließlich Veranstaltungen direkt im Studienseminar haben und somit ein Termin, an dem alle anwesend sind, leicht gefunden werden konnte. Zu späteren Messzeitpunkten konnte dies nicht mehr realisiert werden, da die Ausbildung dann meist in kleineren Gruppen organisiert ist und oftmals direkt an den Ausbildungsschulen stattfindet. Aus diesem Grund wurden zu den folgenden Messzeitpunkten die Fragebogen zum Großteil über die Ausbilderinnen und Ausbilder ausgeteilt. Der Rücklauf der Fragebogen lief über die Sekretariate der Studienseminare. Nach einer Zeitfrist von sechs Wochen wurden die bis dahin zurückgegebenen Fragebogen postalisch an den Autor zurück gesendet.

Alle ReferendarInnen wurden zu Beginn der Untersuchung darüber informiert, dass die Befragung im Rahmen eines Forschungsprojekts durchgeführt wird, welches sich mit Belastungen und Beanspruchungen im Referendariat beschäftigt. Es wurde die vollständig anonyme Auswertung der Daten zugesichert³⁸, die Teilnahme an der Befragung war freiwillig und wurde nicht vergütet. Am Ende der Untersuchung, d.h. nach Realisierung des fünften Messzeitpunkts, wurden alle Teilnehmerinnen über den genauen Untersuchungsanlass und die zentralen Ergebnisse in Form eines schriftlichen Berichts informiert.

7.1.2 Stichprobe

Die Grundgesamtheit, aus der die Stichprobe für die vorliegende Studie gezogen wurde, waren alle ReferendarInnen, die an den 16 kontaktierten Studienseminaren zu t1 mit ihrer Ausbildung begonnen hatten. Potenziell wären zu t1 $N = 503$ ReferendarInnen erreichbar gewesen, von $N = 398$ wurden ausgefüllte Fragebogen zurückgegeben, was einer Rücklaufquote von 79.13% entspricht, die als gut bezeichnet werden kann. In Tabelle 25 sind Anzahl der Befragten und die entsprechenden Rücklaufquoten für die weiteren Messzeitpunkte aufgelistet.

Tabelle 25. Anzahl der Befragten und Rücklaufquoten für t2, t3 und t4 (in Klammern Rücklaufquote in %)

Messzeitpunkt	Anzahl der Befragten	davon wieder zu identifizieren
t1	398	280 ¹
t2	269	230 (57.79) ²
t3	212	146 (63.48) ³
t4	164	94 (64.38) ³

Anmerkungen: 1 = 280 der zum ersten Messzeitpunkt Befragten konnten an mindestens einem der späteren Messzeitpunkte wiederbefragt werden;

2 Die Prozentzahl bezieht sich hier auf die komplette Ausgangsstichprobe am ersten Messzeitpunkt ($N = 398$)

3 Die Prozentzahl bezieht sich auf die Befragten, die zum vorherigen Messzeitpunkt wieder identifiziert werden konnten

Aus Tabelle 25 ist zu ersehen, dass die Rücklaufquoten für t2, t3 und t4 deutlich niedriger liegen als für t1. Dies liegt hauptsächlich darin begründet, dass es für diese Messzeitpunkte

³⁸ Die Zuordnung der jeweiligen Fragebogen eines Befragten erfolgte über einen 8stelligen Code (s. Kap. 7.1.3), der keinen Rückschluss auf die Person ermöglichte.

zumeist nicht möglich war, die ReferendarInnen direkt zu erreichen. In der Mehrzahl der Fälle konnten die Fragebogen nur über die jeweiligen Ausbilderinnen und Ausbilder an den Studienseminaren an die in Frage kommenden ReferendarInnen weiter gegeben werden. Dies spiegelt sich auch in der großen Schwankung der Rücklaufquoten zwischen den Studienseminaren wieder ($M_{\text{über alle Studienseminare und Messzeitpunkte}} = 65\%$; $SD = 21.44$; $\text{Min} = 0\%$; $\text{Max} = 100\%$).

Zur Auswertung des Längsschnitts wurden drei Datensätze gebildet, die im Folgenden beschrieben werden (zur näheren Begründung s. Kap. 7.2.1):

Datensatz „Längsschnitt 1“ (L1) enthält nur diejenigen Befragten, die an allen vier Messzeitpunkten einen Fragebogen ausgefüllt haben. Insgesamt umfasst der Datensatz L1 $N = 94$ ReferendarInnen.

Der zweite Datensatz „Längsschnitt 2“ (L2) beinhaltet diejenigen Befragten, für die von mindestens drei Messzeitpunkten Daten zur Verfügung stehen, wobei die Verteilung der Messzeitpunkte keine Rolle spielt. L2 beinhaltet $N = 190$ ReferendarInnen.

In den dritten Datensatz „Längsschnitt 3“ (L3) gehen die Daten von $N = 300$ ReferendarInnen ein³⁹, für die Angaben von mindestens zwei Messzeitpunkten vorliegen, wobei wiederum die Verteilung der Messzeitpunkte keine Rolle spielt.

In Tabelle 26 ist aufgeführt, aus welchen ReferendarInnen sich die drei Datensätze jeweils zusammensetzen.

Tabelle 26. Beschreibung der drei Datensätze L1, L2, L3 in zentralen soziodemographischen Variablen zum ersten Messzeitpunkt

Datensatz	Alter in Jahren (in Klammern <i>SD</i>)	Geschlecht (davon weiblich)	Partnerschaft (mit PartnerIn)	Kinder (mit einem Kind oder mehreren)
L1	27.66 (4.24)	73.4%	67%	18.1%
L2	27.94 (4.65)	71.6%	65.3%	16.8%
L3	28.38 (4.91)	67.3%	63.7%	17.3%

Tabelle 26 ist zu entnehmen, dass sich die drei gebildeten Datensätze nur unwesentlich in

³⁹ Die Zahl von $N = 300$ kommt deshalb zu Stande, da in diesen Datensatz, wie auch im Datensatz L2, Befragte mit eingehen, die am ersten Messzeitpunkt nicht teilgenommen haben. Insgesamt wurden zu t2 $N = 269$, zu t3 $N = 212$ und zu t4 $N = 165$ ReferendarInnen befragt.

zentralen soziodemographischen Variablen unterscheiden.⁴⁰

7.1.3 Messinstrument

Der Fragebogen für die vier Messzeitpunkte ist im Kern jeweils gleich: Zu allen vier Messzeitpunkten wurden die zentralen Variablen des transaktionalen Stressmodells erhoben, d.h. zu jedem Messzeitpunkt wurde das Ausmaß an sozialer Unterstützung, die Ereigniseinschätzung, die Selbstwirksamkeitserwartungen, Coping, körperliche Beschwerden und psychisches Wohlbefinden erfasst. Das Belastungsausmaß wurde nur zu t2, t3 und t4 erhoben, da zum ersten Messzeitpunkt die ReferendarInnen noch nicht mit den entsprechenden Belastungsfaktoren konfrontiert waren. Neurotizismus als (relativ) stabiles Persönlichkeitsmerkmal wurde nur zu t1 erhoben. Zur Operationalisierung der Variablen wurden die gleichen Skalen wie in Studie 1 (s. Kap. 6.1.3) verwendet⁴¹, auch die Vorgabenreihenfolge war die gleiche wie in Studie 1. Abweichungen zwischen den Messzeitpunkten gab es nur in dem Einsatz von Skalen, die nicht Gegenstand der vorliegenden Arbeit sind. In Tabelle 27 sind die entsprechenden Skalen aufgelistet, im Weiteren wird auf diese Skalen nicht mehr eingegangen.

Tabelle 27. Über die Fragestellung hinausgehende Skalen zu den vier Messzeitpunkten

Messzeitpunkt	Skalen
<i>t1</i>	<ul style="list-style-type: none"> • Identifikation (van Dick et al, 2004) • AVEM (Schaarschmidt & Fischer, 1996)
<i>t2</i>	<ul style="list-style-type: none"> • Identifikation (s.o.)
<i>t3</i>	<ul style="list-style-type: none"> • Identifikation (s.o.) • OCB (van Dick, 1999b)
<i>t4</i>	<ul style="list-style-type: none"> • Identifikation (s.o.) • OCB (s.o.) • AVEM (s.o.)

⁴⁰ Eine statistische Testung der Unterschiede kann in diesem Falle nicht durchgeführt werden, da die Datensätze sich überschneiden, d.h. L1 und L2 sind Substichproben von L3.

⁴¹ Eine Optimierung der Operationalisierung basierend auf den Ergebnissen von Studie 1 war nicht möglich, da der erste Messzeitpunkt parallel zu Studie 1 realisiert wurde und die eingesetzten Skalen zu späteren Messzeitpunkten somit nicht mehr verändert werden konnten, da sonst die Vergleichbarkeit nicht mehr gewährleistet gewesen wäre.

Am Anfang des Fragebogens wurden zunächst soziodemographische Angaben, wie z.B. das Geschlecht, Alter, Familienstand, angestrebtes Lehramt, Größe der Ausbildungsschule, erhoben. Daran anschließend wurden die Komponenten des Stressmodells erfragt. Neben den bereits oben beschriebenen zusätzlichen Skalen wurde auch die Zufriedenheit mit der Anleitung an der jeweiligen Ausbildungsschule der Befragten sowie die Zufriedenheit mit der Ausbildung am Studienseminar erfragt.

Zum Schluss des Fragebogens wurden die ReferendarInnen noch gebeten, einen Code einzutragen, mit dessen Hilfe die Fragebogen eines Befragten der vier Messzeitpunkte einander zugeordnet werden konnten. Dabei sollten die Befragten jeweils den eigenen Geburtstag und den der Mutter eintragen (jeweils nur den Tag und den Monat; z.B. Geburtstag der Mutter am 01. Januar 1911, Geburtstag des Befragten am 12. Dezember 1975: Code ist 01011212). Der Code ist wie die erhobenen Daten generell dritten Personen nicht zugänglich, so dass kein Rückschluss auf die teilnehmenden Personen möglich ist.

Im Anhang ist der Fragebogen für Messzeitpunkt t2 als Beispiel dokumentiert (S. XII), der Bearbeitungsaufwand für alle Fragebogen lag bei etwa 20-30 Minuten.

7.2 *Ergebnisse Studie 2*

Der Ergebnisteil zu Studie 2 untergliedert sich ähnlich wie der Ergebnisteil von Studie 1 in mehrere Abschnitte: Zunächst wird auf das allgemeine Vorgehen bei der Datenanalyse eingegangen (Kap. 1.1.1). Anschließend wird auf die Mortalität eingegangen (Kap. 1.1.2) und die Datenaufbereitung (Kap. 1.1.3) beschrieben. Im nächsten Abschnitt wird das Messmodell analysiert (Kap. 1.1.4) und das Stressmodell zu jedem Messzeitpunkt im Querschnitt geprüft (Kap. 1.1.5), um im Anschluss das Modell einer Überprüfung im Längsschnitt zu unterziehen und mögliche moderierende Einflussvariablen zu untersuchen (Kap. 1.1.6). Abschließend werden die Ergebnisse im Hinblick auf das theoretische Stressmodell diskutiert (Kap. 1.1.7).

7.2.1 Allgemeines Vorgehen bei der Datenanalyse

Um den Zielen der Längsschnittstudie gerecht zu werden, muss zunächst sorgfältig eine adäquate statistische Analyseverfahren ausgewählt werden. Wie bereits formuliert, ist es das Ziel der vorliegenden Längsschnittstudie, die prozessualen Vorhersagen des Stressmodells zu untersuchen. Dabei lassen sich mit Baltes und Nesselroade (1979, S. 21-27) folgende Ziele der Längsschnittuntersuchung formulieren: (a) Die Identifikation von intraindividuellen Veränderungen in den Komponenten des Stressmodells, (b) die Identifikation von interindividuellen

ellen Unterschieden in diesen intraindividuellen Veränderungen, (c) die Analyse von Wechselbeziehungen in diesen Veränderungen; (d) die Analyse von Determinanten intraindividuellere Veränderungen und (e) die Analyse von Determinanten von interindividuellen Unterschieden in intraindividuellen Veränderungen (zur näheren Explikation dieser Ziele in Hinblick auf das Stressmodell siehe Kap. 5). Diesen unterschiedlichen Fragestellungen sollte die auszuwählende Analyseverfahren gerecht werden. Im weiteren wird zunächst auf die klassische Untersuchungsmethode von Längsschnittdaten in den Verhaltenswissenschaften eingegangen, daran anschließend werden neuere Entwicklungen in den Analysemethoden beschrieben, um schließlich eine Methode zur Analyse der vorliegenden Längsschnittdaten vorzustellen.

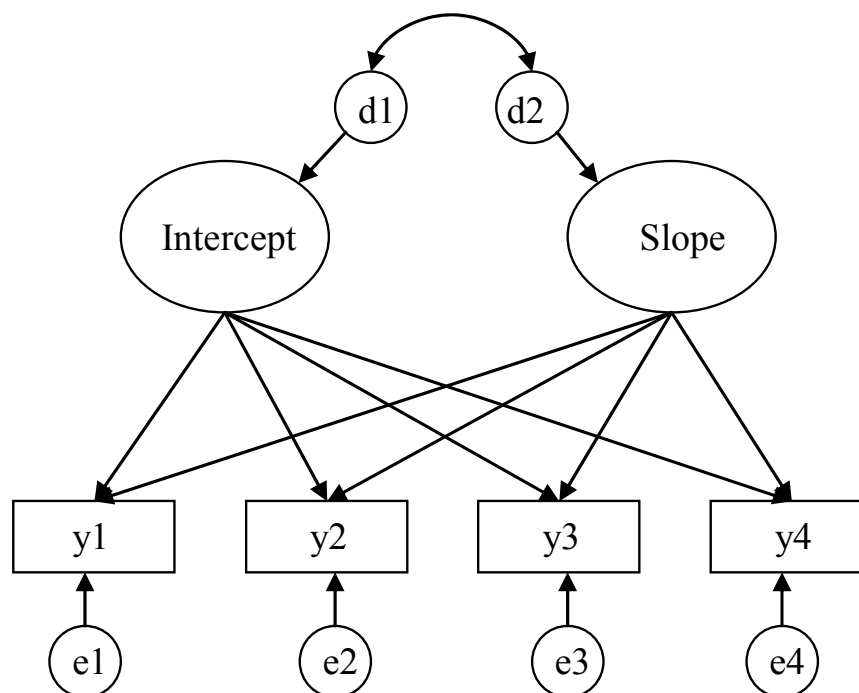
Wie auch in Studie 1 wird zur Analyse der Längsschnittdaten auf Strukturgleichungsmodelle (SEM) zurückgegriffen. Dabei stehen verschiedene Analysetechniken zur Verfügung, eine weit verbreitete Technik stellt das autoregressive cross-lagged Modell dar (ARCL; oder auch „fixed effects Markov simplex model“; Guttman, 1954; Jöreskog, 1979). Zentral für das ARCL-Modell ist, dass spätere Messungen eines Konstrukts durch vorherige Messungen des gleichen Konstrukts vorhergesagt werden (autoregressive Korrelationen). Darüber hinaus können auch frühere Messungen anderer Konstrukte das spätere Konstrukt über die autoregressive Beziehung hinaus vorhersagen (cross-lagged relation). Autoregressive Modelle ermöglichen also zum einen, die Stabilität von Konstrukten über die Zeit abzuschätzen und somit die interindividuelle Stabilität darzustellen, zum anderen aber auch kausale Beziehungen zwischen verschiedenen Konstrukten zu untersuchen. Allerdings wird die Anwendung von ARCL-Modellen kritisch gesehen, wobei sowohl theoretische als auch statistische Aspekte diskutiert werden (Rogosa, 1988, 1995; Urban, 2002; Willett, 1988). Zentral sind drei Einschränkungen (Curran & Hussong, 2002): (a) Das ARCL-Modell ist ein fixed effects Modell, d.h. es wird für alle Regressionsparameter ein Wert geschätzt, der für alle Subjekte einer Stichprobe gleich ist. Der Einfluss wird somit für alle Subjekte gleich gesetzt, was oftmals eine unrealistische Bedingung ist und gerade auch im Hinblick für die vorliegenden Fragestellungen kritisch zu sehen ist; (b) die individuelle Mittelwertsstruktur ist in solchen Modellen nicht beinhaltet, intraindividuelle Mittelwertsveränderungen können somit nicht dargestellt werden; (c) die Veränderungen in einem Konstrukt zwischen zwei Messzeitpunkten ist unabhängig von dem Einfluss früherer Veränderungen und späterer Änderungen. Mehrere Wiederholungsmessungen werden damit in eine Serie von Vergleichen zwischen zwei Messzeitpunkten zerlegt, was meist weder mit der Theorie noch mit der Struktur der beobachteten empirischen Daten übereinstimmt. Gerade für Fragestellungen, die (auch) die

individuellen Unterschiede in Entwicklungsverläufen beinhalten, sind ARCL-Modelle nicht die Methode der Wahl, da interindividuelle Unterschiede in den Veränderungen von Merkmalen nicht dargestellt werden können.

Unter anderem hat die Kritik an ARCL-Modellen zu der Entwicklung alternativer statistischer Modelle geführt (für einen Überblick s. z.B. Hertzog & Nesselroade, 2003; Moskowitz & Hershberger, 2002). Für die vorliegende Studie von besonderer Relevanz stellen latente Wachstumskurvenmodelle (englisch „latent growth curve models“; LGC) dar (Browne & Du Toit, 1991; Duncan & Duncan, 1995; McArdle, 1986, 1988, 1998; Meredith & Tisak, 1990; Willett & Sayer, 1994). Die grundlegenden Aspekte von LGC sind in Abbildung 6 dargestellt. In LGC-Modellen wird basierend auf den Wiederholungsmessungen eine einzelne Trajektorie⁴² für jede Person über alle Messzeitpunkte hinweg geschätzt, die Kurve besteht aus einem latenten Intercept-Faktor und einem latenten Slope-(Steigungs-)Faktor. Die Intercepts und Slopes aller Personen werden gemittelt und stellen dann den mittleren Intercept und mittleren Slope der untersuchten Stichprobe dar (fixed effects). Darüber hinaus wird die Variabilität in dem mittleren Intercept- und Slope-Faktor geschätzt (random effects). LGC-Modelle modellieren somit nicht nur die beobachtete Kovarianz-Struktur der Daten, sondern auch die Mittelwerts-Struktur.

⁴² Eine Trajektorie (deutsch *Bahnkurve*) bezeichnet im Allgemeinen die Flugbahn eines Objektes. Im engeren Sinne handelt es sich aber um die "Entwicklungslinie" eines dynamischen Systems im Phasenraum. Auf den vorliegenden Kontext angewendet stellt eine Trajektorie eine kontinuierliche Veränderung in einem Merkmal über die Zeit dar.

Abbildung 6. Ein univariates LGC-Modell für eine Variable, die zu vier Messzeitpunkten erhoben wurde (y1 bis y4)



Das Beispiel in Abbildung 6 zeigt, dass für eine Variable, welche zu mehreren Messzeitpunkten erhoben wird, zwei latente Faktoren geschätzt werden, ein latenter Intercept-Faktor und ein latenter Slope-Faktor. Der Intercept-Faktor stellt den mittleren Ausgangswert aller Personen dar, der Slope-Faktor die mittlere Steigung der Trajektorie (also die Veränderung in dem mittleren Wert über die Messzeitpunkte hinweg). Die Varianz in den Intercept- und Slope-Faktoren repräsentiert die individuelle Variabilität in den mittleren Werten. Schließlich können die beiden latenten Faktoren interkorrelieren, was inhaltlich bedeutet, dass die Veränderung in einem Merkmal abhängig ist von dem Ausgangsniveau in demselbigen.

LGC-Modelle können beliebig erweitert werden um weitere Variablen. Beispielsweise können einem LGC-Modell weitere Variablen hinzugefügt werden, die den Intercept und den Slope beeinflussen können. Letztlich können LGC-Modelle die meisten der oben genannten Ziele von Längsschnittuntersuchungen bedienen (McArdle & Bell, 2000), es können (a) intraindividuelle Veränderungen direkt geschätzt werden (Varianz des Slope-Faktors); (b) interindividuelle Unterschiede in den intraindividuellen Veränderungen geschätzt werden (z.B. durch Hinzunahme von weiteren Variablen oder durch Vergleich von Subgruppen); (c) wechselseitige Abhängigkeiten in den Veränderungen mehrerer Variablen bestimmt werden; (d) Determinanten für intraindividuelle Veränderungen in LGC-Modellen auf-

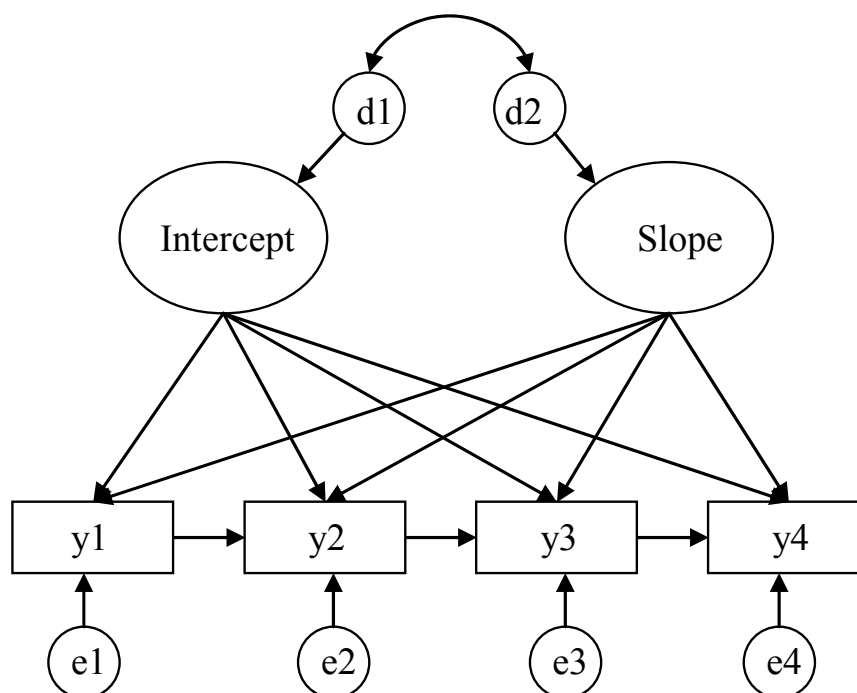
genommen werden und schließlich (e) die Ursachen von interindividuellen Unterschieden in intraindividuellen Veränderungen untersucht werden.

Gerade die Punkte (a), (b), (d) und (e) sind zur Beantwortung der Hypothesen der vorliegenden Untersuchung von besonderer Bedeutung, da (1) untersucht werden soll, ob im Verlauf des Referendariats die körperlichen Beschwerden zunehmen und ob sich das psychische Wohlbefinden verschlechtert, (2) ob diese Veränderungen von proximalen (Coping) und distalen Prädiktoren (Ereigniseinschätzung und Selbstwirksamkeitserwartungen) abhängig sind und (3) ob die Prädiktoren und die Veränderungen durch Personen- und Umweltmerkmale (Neurotizismus und Ausmaß an sozialer Unterstützung) beeinflusst werden. Somit erscheint die Verwendung von LGC-Modellen zur Auswertung der Längsschnittdaten der vorliegenden Studie indiziert.

Allerdings soll auch die kausale Beziehung der Modellvariablen untersucht werden. Es ist somit auch von Interesse, inwiefern die Modellvariablen über die Zeit miteinander zusammenhängen, ideal wäre somit die Modellierung von cross-lagged-Effekten. Auch die Stabilität der Merkmale über die Zeit stellen wichtige Informationen dar (lagged-Effekte). Reine LGC-Modelle bilden sowohl cross-lagged- als auch lagged-Effekte nicht ab. Das ARCL- und LGC-Modell sind somit jeweils nur für bestimmte Formen von zeitlichen Veränderungen geeignet. Das ARCL-Modell modelliert explizit zeitspezifische Beziehungen innerhalb und zwischen Wiederholungsmessungen von einem oder mehreren Konstrukten, während das LGC-Modell diese Beziehungen nur auf der Ebene von kontinuierlichen latenten Trajektorien modelliert, die den Wiederholungsmessungen unterliegen. Gerade auch bezogen auf die vorliegende Untersuchung wäre es wünschenswert, beide Modelle miteinander zu kombinieren, um so die jeweiligen Stärken der beiden Modelle zu nutzen. Gleichzeitig würde eine solche Kombination auch die Entscheidung ermöglichen, ob ein ARCL-, ein LGC-Modell oder eine Kombination beider Modelle die Daten am besten beschreiben.

Bollen und Curran (2004; Curran & Bollen, 2001; für ein Anwendungsbeispiel s. Davydov, Schmidt & Bamberg, 2004) haben jüngst ein so genanntes Autoregressive-Latent-Trajectory (ALT)-Modell vorgestellt, welches sowohl das LGC- als auch das ARCL-Modell subsumiert (s. Abbildung 7).

Abbildung 7. Ein univariates ALT-Modell für eine Variable, die zu vier Messzeitpunkten erhoben wurde (y1 bis y4)



Für die vorliegende Untersuchung ist das ALT-Modell optimal, um die offenen Fragen beantworten zu können. Das genaue Vorgehen wird in Kapitel 7.2.6 beschrieben.

Wie bereits im Methodenteil beschrieben, wurden drei unterschiedliche Datensätze gebildet: L1, L2 und L3. L1 umfasst nur solche ReferendarInnen, die zu allen vier Messzeitpunkten einen Fragebogen bearbeitet haben ($N=94$), L2 beinhaltet alle Befragten, die an mindestens drei Messzeitpunkten teilgenommen haben ($N=190$), L3 berücksichtigt auch solche Personen, für die nur von zwei Messzeitpunkten Daten vorliegen ($N=300$). Dieser Weg wurde gewählt, da der Anteil derjenigen Befragten, die an allen Messzeitpunkten teilgenommen haben, sehr klein ist und somit zum einen bei komplexen Modellen die statistische Power gering (MacCallum et al. 1996; MacCallum & Hong, 1997) und zum anderen die Güte der Parameterschätzung gefährdet ist (Bentler & Yuan, 1999; Jackson, 2001, 2003). Aus diesem Grund wird zur Modellberechnung der Datensatz L2 verwendet, um das Modell zu prüfen. Die Datensätze L1 und L3 dienen dazu, das so gewonnene Modell auf seine Stabilität hin zu prüfen.

Zur Abschätzung der Güte eines Modells wird wie in Studie 1 (s. Kap. 6.2.1) auf mehrere Fit-Indizes zurückgegriffen, wobei dem Vorschlag von Hu und Bentler (1999) gefolgt wird: Der Fit eines Modells wird dann als gut bewertet, wenn der CFI $\approx .95$ und der RMSEA $\approx .06$

beträgt.⁴³ Zur Vollständigkeit wird die χ^2 -goodness-of-fit-Statistik sowie das Verhältnis von χ^2/df angegeben, welches idealerweise 2:1 sein sollte. Darüber hinaus wird wiederum für alle Endmodelle (Messmodelle, Strukturmodell) der p-close-Fit angegeben, der größer .05, idealerweise größer .10 sein sollte (vgl. Schermelleh-Engel et al., 2003).

Als statistisches Verfahren zur Modellschätzung und Modellbewertung wird das Maximum-Likelihood-Schätzverfahren (ML) verwendet (Browne, 1982; Jöreskog, 1967), als Input dienen Kovarianzmatrizen.

7.2.2 Panel-Mortalität

In diesem Abschnitt soll geprüft werden, ob eine systematische Panel-Mortalität vorliegt, die zu einer Verzerrung der Ergebnisse von Studie 2 führen könnte (Little, Lindenberger & Maier, 2000). Eine selektive Mortalität, d.h. die Nicht-Teilnahme an den weiteren Befragungen (Dropout), ist durch bestimmte Merkmale der Befragten verursacht und führt dazu, dass die Ergebnisse nicht auf die ursprüngliche Stichprobe bzw. Population generalisiert werden können. Gleichzeitig besteht aber auch die Möglichkeit, dass die Befragten, die weiterhin an der Längsschnittuntersuchung teilnehmen (Continuers)⁴⁴, bestimmte Merkmale aufweisen, die letztlich zu den ermittelten Ergebnissen geführt haben.

Nach Little et al. (2000; s.a. Lindenberger et al., 1999) sind zwei Arten von Selektivitätsfragen zu unterscheiden: (1) Eine Form von Selektivität bezieht sich auf multivariate Unterschiede zwischen Dropouts und Continuers. Dabei wird geprüft, ob es in den gemessenen Variablen und Konstrukten systematische Unterschiede zwischen diesen Gruppen gibt. Hierzu können latente Multigruppen-Analysen herangezogen werden, ein Hinweis auf selektiven Dropout liegt vor, wenn es zwischen den Dropouts und den Continuers signifikante Unterschiede in den Mittelwerten, Varianzen und/oder Kovarianzen gibt. (2) Die zweite Form von Selektivität bezieht sich auf den vorhergesagten Effekt unterschiedlicher Teilnahmeprofile, wenn unterschiedliche Ebenen von Teilnahme berücksichtigt werden. Anders formuliert wird hierbei der Frage nachgegangen, ob die Beziehungen zwischen den Konstrukten anders geartet wären, wenn die Dropouts in die Endstichprobe mit eingegangen wären. Zur Beantwortung dieser Fragestellung gibt es verschiedene Analysetechniken (für einen Über-

⁴³ Der SRMR kann nicht zur Bewertung der Modelle herangezogen werden, da bei Verwendung von FIML (s.u.) der SRMR nicht berechnet werden kann.

⁴⁴ Unter Continuers fallen alle Befragten, die an mindestens einem weiteren Messzeitpunkt teilgenommen haben und somit zumindest einem der drei gebildeten Längsschnittdatensätzen (L1, L2, L3) zugeordnet sind.

blick s. Little et al., 2000), in der vorliegenden Studie wird auf den FIML- Ansatz zurückgegriffen (s. weiter u.).

Neben diesen beiden Fragestellungen wird zusätzlich geprüft, ob sich die Dropouts von den Continuers in wichtigen soziodemographischen Variablen unterscheiden.⁴⁵ Fragestellung 1, also die Frage nach Unterschieden in den Mittelwerten, Varianzen und Kovarianzen wird mit einer Multigruppenanalyse überprüft. Dagegen wird Fragestellung 2 erst zu einem späteren Zeitpunkt geklärt, wenn es um die Überprüfung des Messmodells und des ALT-Modells geht (s. Kapitel 7.2.4 und 7.2.6.), da an dieser Stelle die FIML-Methode zum Einsatz kommt. Zunächst soll aber der Frage nach Unterschieden in wichtigen soziodemographischen Variablen nachgegangen werden.

Der Vergleich der Dropouts ($N = 118$) mit den Continuers ($N = 280$) zeigt (s. Tabelle 28), dass es einen signifikanten Unterschied im durchschnittlichen Alter zwischen beiden Gruppen gibt (Dropouts: $M = 28.40$, $SD = 4.92$; Continuers: $M = 29.86$, $SD = 5.64$; $t(396) = 2.58$, $p < .01$). Die Befragten der Dropout-Gruppe sind im Durchschnitt älter als die Befragten der Continuers-Gruppe. In allen weiteren relevanten Merkmalen unterscheiden sich die beiden Gruppen nicht voneinander.

⁴⁵ Zwar ist dies der übliche Ansatz, um eine systematische Selektivität in Längsschnittuntersuchungen zu prüfen. Allerdings stellen die oben dargestellten Selektivitätsfragen wichtige Erweiterungen dar, die mit dem klassischen Vorgehen nicht beantwortet werden können (s. Little et al., 2000).

Tabelle 28. Verteilung der Befragten auf Geschlecht, Studienseminar, Familienstand und Kinder (in Klammern in Prozent) getrennt nach Dropouts und Continuers

		Dropouts	Continuers	χ^2 -Test
		(N = 118)	(N = 280)	
<i>Geschlecht</i>	Männlich	37 (31.4)	80 (28.6)	$\chi^2 = .31$, $p = .63$
	Weiblich	81 (68.6)	200 (71.4)	
<i>Studienseminar</i>	Berufliche Schulen	13 (11)	38 (13.6)	$\chi^2(2) = 1.72$, $p = .42$
	GHRs	58 (49.2)	149 (53.2)	
	Gymnasien	47 (39.8)	93 (33.2)	
<i>Familienstand</i>	Ledig	42 (35.6)	90 (32.1)	$\chi^2 = .45$, $p = .56$
	Partner/Partnerin	76 (64.4)	190 (67.9)	
<i>Kinder</i>	Nein	94 (79.7)	223 (79.6)	$\chi^2 = .07$, $p = .78$
	Ja	24 (20.3)	53 (18.9)	

Sowohl von den Dropouts als auch den Continuers stehen zum ersten Messzeitpunkt neben den oben aufgelisteten soziodemographischen Variablen noch Daten zu Neurotizismus, psychischem Wohlbefinden und körperlichen Beschwerden zur Verfügung.⁴⁶ Mittels einer Multigruppenanalyse (Bollen, 1989; Byrne, 2001; Kline, 1998) wurde überprüft, ob sich die beiden Gruppen in den Mittelwerten, Varianzen und Kovarianzen dieser Variablen unterscheiden. Hierzu wurden mehrere hierarchisch abgestufte Modelle berechnet, wobei ausgehend von einem Modell ohne Restriktion die Modelle immer restriktiver werden:

- *Modell 1.* ohne Restriktion (Alle Parameter können in beiden Gruppen variieren, Basismodell)
- *Modell 2.* Faktorladungen sind in beiden Gruppen gleich
- *Modell 3.* wie 2 und Messintercepts sind gleich
- *Modell 4.* wie 3 und Interkorrelationen sind gleich
- *Modell 5.* wie 4 und Fehlervarianzen sind gleich

⁴⁶ Die genaue Zusammensetzung der latenten Faktoren, d.h. welche Indikatoren jeweils für Neurotizismus, psychisches Wohlbefinden und körperliche Beschwerden herangezogen wurden, wird ausführlich im Anhang auf S. XXV dargestellt.

Alle Modelle zeigen eine gute Datenanpassung, die Fit-Indizes unterscheiden sich zwischen den Modellen nur marginal (s. Tabelle 29).

Tabelle 29. Fit-Indizes der Modelle der Multigruppenanalyse zur Überprüfung der Übereinstimmung in den Mittelwerten, Varianzen und Kovarianzen zwischen den Dropouts und den Continuern

Modell	χ^2	df	<i>p</i>	CFI	RMSEA	χ^2/df
1	101.92	48	<.001	.97	.05	2.12
2	103.59	54	<.001	.97	.05	1.92
3	111.73	63	<.001	.97	.04	1.77
4	114.97	69	<.001	.97	.04	1.67
5	125.18	78	<.001	.97	.04	1.61

Anmerkungen: *Modell 1.* ohne Restriktion; kein Parameter wurde zwischen beiden Stichproben gleichgesetzt; *Modell 2.* Faktorladungen sind in beiden Gruppen gleich; *Modell 3.* wie 2 und Messintercepts sind gleich; *Modell 4.* wie 3 und Interkorrelationen sind gleich; *Modell 5* wie 4 und Fehlervarianzen sind gleich

Ein genesteter Modellvergleich zeigt, dass sich bei Annahme, dass Modell 1 gültig ist, Modell 2 keine signifikant schlechtere Datenanpassung aufweist ($\Delta\chi^2 = 1.66$, $df = 6$, $p = .95$). Ebenso zeigt Modell 3 bei Annahme, dass Modell 2 gültig ist, keine schlechtere Datenanpassung ($\Delta\chi^2 = 8.14$, $df = 9$, $p = .52$). Auch Modell 4 im Vergleich zu Modell 3 ($\Delta\chi^2 = 3.24$, $df = 6$, $p = .78$) und Modell 5 im Vergleich zu Modell 4 ($\Delta\chi^2 = 10.22$, $df = 9$, $p = .33$) repräsentieren die Daten gleich gut. Somit zeigt sich, dass Dropouts und Continuers hinsichtlich der Mittelwerte, Varianzen und Kovarianzen in Neurotizismus, psychischem Wohlbefinden und körperlichen Beschwerden vergleichbar sind.

Basierend auf den beiden präsentierten Analysen ist nicht endgültig zu entscheiden, ob sich Dropouts und Continuers systematisch unterscheiden. Allerdings deuten die bisherigen Analysen darauf hin, dass von keiner systematischen Mortalität auszugehen ist. Abschließend lässt sich dies allerdings erst konstatieren, wenn die oben angesprochenen FIML-Analysen durchgeführt wurden.

7.2.3 Datenaufbereitung

Wie bereits weiter oben erwähnt, wurden insgesamt drei Datensätze gebildet. Sowohl fehlende Werte in einzelnen Items als auch das komplette Fehlen von Daten einer Versuchsperson

son zu einem oder maximal zwei Messzeitpunkten (relevant für die Datensätze L2 und L3) werden mit FIML (s. Abschnitt X) geschätzt. Bislang liegen keine Simulationsstudien vor, die untersucht haben, ab welchem Ausmaß fehlender Werte die FIML-Methode zu verzerrten Ergebnissen führt. Wothke (2000) verwendet in einer Simulationsstudie Datensätze mit bis zu 40% fehlender Werte zu einem Messzeitpunkt und kann zeigen, dass FIML im Vergleich zu anderen Methoden zum Umgang mit fehlenden Werten zu den besten Ergebnissen führt.

Auch für die vorliegende Studie wurde analysiert, wie hoch der maximale prozentuale Ausfall zu den einzelnen Messzeitpunkten ist. Der Datensatz L1 setzt sich wie beschrieben aus denjenigen Befragten zusammen, die an allen Messzeitpunkten teilgenommen haben. Der maximale Anteil fehlender Werte in den zentralen Modellvariablen liegt bei maximal drei Prozent zu allen Messzeitpunkten, für den Großteil der Variablen liegt der Anteil fehlender Werte bei weniger als einem Prozent. In Tabelle 30 sind für die Datensätze L2 und L3 der prozentuale Anteil an Dropouts⁴⁷ für jeden Messzeitpunkt aufgelistet.

Tabelle 30. Anteil an Dropouts (in Klammern in Prozent) zu allen Messzeitpunkten für die Datensätze L2 und L3

Messzeitpunkt	Datensatz	
	L2 (N = 190)	L3 (N = 300)
t1	4 (2.1)	20 (6.7)
t2	20 (10.5)	54 (18)
t3	22 (11.6)	102 (34)
t4	51 (26.8)	141 (47)

Tabelle 30 ist zu entnehmen, dass die Ausfallquote in dem Datensatz L2 moderat ist, im Datensatz L3 dagegen zumindest zum vierten Messzeitpunkt substantiell ist. Dies stützt gleichzeitig das Vorgehen, zunächst mit dem Datensatz L2 alle Analysen anzugehen, da hier davon ausgegangen werden kann, dass die auf der FIML-Methode basierenden Parameterschätzer effizient sind (s. Wothke, 2000).

⁴⁷ Der Begriff Dropouts wird auch für solche Fälle verwendet, die erst zu einem späteren Messzeitpunkt an der Befragung teilgenommen haben oder die einen (bzw. maximal zwei) Messzeitpunkte ausgelassen haben, dann aber wieder an der Befragung teilgenommen haben (so genannte Returner; Little et al., 2000). Aus diesem Grund gibt es auch für den ersten Messzeitpunkt Dropouts.

Neben dem Ausmaß fehlender Werte wurde zusätzlich überprüft, ob die Daten multivariat normalverteilt sind, da dies eine Voraussetzung für Strukturgleichungsanalysen ist (West et al., 1995; s. Kap. 6.2.3.3). Die Prüfung auf Normalverteilung erfolgte in allen drei Datensätzen zu allen Messzeitpunkten nur in den zentralen Konstruktvariablen. Da in Studie 2 auf das gleiche Messmodell wie in Studie 1 zurückgegriffen wurde (s.u.), wurde die Prüfung auf Parcel-Ebene vorgenommen. Die Prüfung auf Normalverteilung erfolgte mit AMOS 5.0. Bei Verwendung der FIML-Methode kann die Normalverteilungsprüfung nicht vorgenommen werden, aus diesem Grund wurden für die Datensätze L2 und L3 fehlende Werte in den Konstrukt-Parceln mit der EM-Methode ersetzt (Enders, 2001; s. Kap. 6.2.3.2), um auch für diese Datensätze eine Normalverteilungsprüfung vornehmen zu können. Tabelle 31 fasst die Ergebnisse für alle drei Datensätze zusammen.

Tabelle 31. Mittlere Schiefe und Kurtosis (in Klammern Standardabweichung) in den zentralen Modellvariablen sowie Mardia-Koeffizient zu allen Messzeitpunkten getrennt für die Datensätze L1, L2 und L3

	Datensatz		
	L1	L2	L3
<i>Messzeitpunkt 1</i>			
Kurtosis	-0.03 (0.59)	0.35 (0.90)	0.33 (0.85)
Schiefe	0.12 (0.62)	0.17 (0.79)	0.13 (0.79)
Mardia-Koeffizient	6.69*	10.34***	14.33***
<i>Messzeitpunkt 2</i>			
Kurtosis	-0.08 (0.45)	0.30 (0.79)	0.20 (0.52)
Schiefe	-0.09 (0.59)	-0.12 (0.71)	-0.10 (0.64)
Mardia-Koeffizient	31.34***	52.05***	77.81***
<i>Messzeitpunkt 3</i>			
Kurtosis	0.16 (0.70)	0.24 (0.73)	0.24 (0.56)
Schiefe	-0.06 (0.68)	-0.08 (0.70)	-0.03 (0.58)
Mardia-Koeffizient	47.64***	54.82***	89.31***
<i>Messzeitpunkt 4</i>			
Kurtosis	-0.06 (0.68)	0.04 (0.54)	0.19 (0.44)
Schiefe	-0.16 (0.60)	-0.10 (0.58)	-0.07 (0.49)
Mardia-Koeffizient	34.35***	48.09***	87.41***

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

Die mittlere Kurtosis und mittlere Schiefe ist für alle Datensätze und Messzeitpunkte in einem akzeptablen Bereich (West et al., 1995), wenn auch der Mardia-Koeffizient jeweils signifikant wird. Somit kann von einer geringen Abweichung von der multivariaten Normalverteilung ausgegangen werden. Wie bereits in Studie 1 erwähnt (s. Kap. 6.2.3.3; s.a. Olsson et al., 2000), ist die gewählte ML-Schätzfunktion relativ robust gegenüber kleineren Abweichungen von der multivariaten Normalverteilung. Alle finiten Modelle (Messmodelle als auch ALT-Modell) werden mit der Bootstrap-Methode gegen geprüft (Efron, 1979, 1982; näheres zur Bootstrap-Methode findet sich in Kap. 6.2.3.3), um eventuelle Verzerrungen, verursacht durch die Abweichung von der Normalverteilung, zu analysieren. Allerdings ist dies in AMOS nur möglich mit Datensätzen ohne fehlende Werte. Aus diesem Grund wird auch hier auf die modifizierten Datensätze L1, L2 und L3 zurückgegriffen, in denen die fehlenden Werte mit der EM-Methode ersetzt wurden, um die Bootstrap-Methode zu ermöglichen. Da Ziel dieser Analyse ist, mögliche Verzerrungen in den Parameterschätzern abzuschätzen, die durch die Abweichung von der Normalverteilung verursacht worden sind, scheint dieses Vorgehen gerechtfertigt.

7.2.4 Überprüfung des Messmodells

Die Überprüfung des Messmodells jeweils für die drei Datensätze erfolgt mehrschrittig. Zunächst wird jeweils für alle Datensätze zu jedem Messzeitpunkt das Messmodell geprüft. Daran anschließend wird analysiert, ob das Messmodell über die Messzeitpunkte hinweg stabil bleibt. Abschließend werden die Modelle erneut mit der Bootstrap-Methode berechnet, um eventuelle Verzerrungen durch die oben konstatierte Abweichung von der Normalverteilung abzuschätzen.

Dieses Vorgehen wird gewählt, da in Längsschnittstudien nur dann die Messzeitpunkte aufeinander bezogen werden können, wenn die Messmodelle zu allen Zeitpunkten miteinander vergleichbar sind (Meredith & Horn, 2001; Pitts, West & Tein, 1996; Vandenberg & Lance, 2000). Es gibt verschiedene Formen der Invarianz von Messmodellen (Steinmetz et al., in press; Widaman, 2000), auf die bereits in Kapitel 6.2.5.2 hingewiesen wurde und die generell auch auf Längsschnittdaten angewendet werden müssen. Für Längsschnittanalysen am bedeutsamsten sind die *konfigurale* und *metrische* Invarianz sowie Invarianz der Faktor-Kovarianzen bzw. Interkorrelation (Meredith & Horn, 2001; Vandenberg & Lance, 2000). *Konfigurale* Invarianz liegt vor, wenn zu allen Messzeitpunkten die gleiche Anzahl an Faktoren vorliegt und das Muster von fixierten und freien Faktorladungen gleich ist. Von *metri-*

scher Invarianz spricht man, wenn über alle Messzeitpunkte die Faktorladungen gleich sind. Bei vorliegender metrischer Invarianz sind die Faktoren inhaltlich zwischen den verschiedenen Messzeitpunkten zu vergleichen. Diese Form von Invarianz muss also vorliegen. Tisak und Meredith (1991) bezeichnen dies auch als *stationarity*⁴⁸ (s.a. Meredith & Horn, 2001; Pitts et al., 1996). Invarianz der Faktor-Kovarianzen meint, dass über alle Messzeitpunkte die Interkorrelation zwischen den latenten Faktoren gleich ist.

Neben diesen drei Formen von Invarianz gibt es noch weitere Formen von Invarianz wie z.B. skalare Invarianz, die Invarianz der Faktorvarianzen und latenten Mittelwerte sowie der Fehlervarianzen. Diese sind aber für Längsschnittuntersuchungen von geringerer Bedeutung bzw. z.T. kontraindiziert, da beispielsweise Mittelwertsveränderungen über die Zeit erwartet werden. Aus diesem Grund werden in Studie 2 nur die oben beschriebenen Formen von Invarianz für das Messmodell überprüft.

Das zu prüfende Messmodell entspricht in Großteilen dem aus Studie 1 (s. Kap. 6.2.5). Als Indikatoren für die jeweiligen Konstrukte wurde auf die gleichen Parcels zurückgegriffen (zur Entwicklung und Zusammensetzung der Parcels s. Kap. 6.2.4.2). Zu beachten sind folgende Abweichungen⁴⁹: (1) Zum ersten Messzeitpunkt wurden nur Angaben zu Neurotizismus, psychischem Wohlbefinden und körperlichen Beschwerden erhoben. Somit wird hier nur ein stark reduziertes Modell geprüft. (2) Zu allen weiteren Messzeitpunkten (Messzeitpunkte zwei, drei und vier) wurde auf eine Erhebung von Neurotizismus verzichtet, somit enthalten zu diesen Messzeitpunkten die Modelle keinen Neurotizismusfaktor.

In Tabelle 32 sind die Ergebnisse der CFAs für alle Messzeitpunkte aufgeführt, jeweils getrennt für die Datensätze L1, L2 und L3. Faktorladungen und Interkorrelationen sind im Anhang dokumentiert (s. S. XXVIff; Tabellen A9-A20).

⁴⁸ Der Begriff kann nur schwer ins Deutsche übersetzt werden. Gemeint ist Unveränderlichkeit.

⁴⁹ Diese Abweichungen bedingen auch, dass nicht mit einer Multigruppenanalyse überprüft werden kann, ob es Abweichungen in dem Messmodell zwischen Studie 1 und Studie 2 gibt.

Tabelle 32. Ergebnisse der CFAs zur Überprüfung der Messmodelle zu allen Messzeitpunkten getrennt nach Datensätzen

Messzeitpunkt	Datensatz		
	L1	L2	L3
t1	$\chi^2 = 34.27$, df = 24, $p = .08$, CFI = .96, RMSEA = .07, p-close = .27, $\chi^2/\text{df} = 1.43$	$\chi^2 = 54.63$, df = 24, $p < .001$, CFI = .95, RMSEA = .08, p-close = .04, $\chi^2/\text{df} = 2.28$	$\chi^2 = 61.14$, df = 24, $p < .001$, CFI = .97, RMSEA = .07, p-close = .05, $\chi^2/\text{df} = 2.55$
t2	$\chi^2 = 199.92$, df = 137, $p < .001$, CFI = .92, RMSEA = .07, p-close = .07, $\chi^2/\text{df} = 1.46$	$\chi^2 = 266.30$, df = 137, $p < .001$, CFI = .92, RMSEA = .07, p-close = .01, $\chi^2/\text{df} = 1.94$	$\chi^2 = 282.92$, df = 137, $p < .001$, CFI = .94, RMSEA = .06, p-close = .05, $\chi^2/\text{df} = 2.07$
t3	$\chi^2 = 201.87$, df = 137, $p < .001$, CFI = .93, RMSEA = .07, p-close = .06, $\chi^2/\text{df} = 1.47$	$\chi^2 = 274.26$, df = 137, $p < .001$, CFI = .92, RMSEA = .07, p-close < .01, $\chi^2/\text{df} = 2.00$	$\chi^2 = 254.12$, df = 137, $p < .001$, CFI = .94, RMSEA = .05, p-close = .28, $\chi^2/\text{df} = 1.86$
t4	$\chi^2 = 201.16$, df = 137, $p < .001$, CFI = .93, RMSEA = .07, p-close = .06, $\chi^2/\text{df} = 1.47$	$\chi^2 = 213.19$, df = 137, $p < .001$, CFI = .95, RMSEA = .05, p-close = .30, $\chi^2/\text{df} = 1.56$	$\chi^2 = 237.07$, df = 137, $p < .001$, CFI = .94, RMSEA = .05, p-close = .52, $\chi^2/\text{df} = 1.73$

Es lässt sich zusammenfassen, dass zu allen Messzeitpunkten in allen Datensätzen die Messmodelle ausreichend an die Daten angepasst sind, wenn auch teilweise die Fit-Werte nicht ganz die gewünschten Werte erreichen. Dennoch scheint sich die in Studie 1 vorgenommene Bildung der latenten Faktoren bewährt zu haben.

Wie bereits oben erläutert, ist für die vorliegende Studie noch bedeutsamer, ob das Messmodell über die Messzeitpunkte vergleichbar ist, zumindest stationarity sollte gegeben sein. Um dies zu prüfen, wurden, jeweils getrennt für jeden Datensatz, zwei Multigruppenanalysen durchgeführt, um die Messmodelle zu den verschiedenen Messzeitpunkten zu vergleichen. Es müssen zwei Analysen durchgeführt werden, da zum ersten Messzeitpunkt nur Neurotizismus, körperliche Beschwerden und psychisches Wohlbefinden erhoben wurden und somit ein direkter Vergleich mit den Messmodellen zu den übrigen Messzeitpunkten nicht erfolgen konnte. Aus diesem Grund wurden in der *Multigruppenanalyse 1* nur die Messmodelle für die Messzeitpunkte zwei, drei und vier miteinander verglichen, in der *Multigruppenanalyse 2* wurde, um die stationarity für psychisches Wohlbefinden und körperliche Beschwerden einschließlich dem ersten Messzeitpunkt zu prüfen, ein reduziertes Modell berechnet, welches nur die latenten Faktoren von psychischem Wohlbefinden und körperlichen Beschwerden umfasst.

Für Multigruppenanalyse 1 wurden mehrere hierarchisch abgestufte Modelle berechnet, wobei ausgehend von einem Modell ohne Restriktion die Modelle immer restriktiver werden:

- *Modell 1.* ohne Restriktion (Alle Parameter können über die Messzeitpunkte variieren, Basismodell; Test auf konfigurale Invarianz)
- *Modell 2.* Faktorladungen sind zu allen Messzeitpunkten gleich (Test auf metrische Invarianz bzw. stationarity)
- *Modell 3.* wie 2 und Interkorrelationen sind gleich (Test auf Invarianz der Faktor-Kovarianzen)

Die Ergebnisse zeigen zunächst, dass in allen drei Datensätzen Modell 1 sehr gut an die Daten angepasst ist (s. Tabelle 33), somit kann von konfiguraler Invarianz der Messmodelle zu den Messzeitpunkten t2, t3 und t4 ausgegangen werden.

Tabelle 33. Ergebnisse für Modell 1 (Test auf konfigurale Invarianz) der Multigruppenanalyse 1 für alle Datensätze

	Datensatz		
	L1	L2	L3
Modell 1	$\chi^2 = 602.95$, df = 411, $p < .001$, CFI = .93, RMSEA = .04, p-close = .99, $\chi^2/\text{df} = 1.47$	$\chi^2 = 753.75$, df = 411, $p < .001$, CFI = .93, RMSEA = .04, p-close = 1.00, $\chi^2/\text{df} = 1.83$	$\chi^2 = 774.11$, df = 411, $p < .001$, CFI = .94, RMSEA = .03, p-close = 1.00, $\chi^2/\text{df} = 1.88$

Anmerkung: *Modell 1.* ohne Restriktion, kein Parameter wurde zwischen den Messzeitpunkten gleichgesetzt

Im Folgenden wird nur auf die Ergebnisse für den Datensatz L2 eingegangen, für die Datensätze L1 und L3 werden gleiche Ergebnisse ermittelt.

Der genestete Modellvergleich zeigt, dass sich bei Annahme, dass Modell 1 gültig ist, Modell 2 keine signifikant schlechtere Datenanpassung aufweist ($\Delta\chi^2 = 20.17$, df = 26, $p = .78$). Ebenso zeigt Modell 3 bei Annahme, dass Modell 2 gültig ist, keine schlechtere Datenanpassung ($\Delta\chi^2 = 37.65$, df = 42, $p = .66$).

Die Ergebnisse der Multigruppenanalyse 1 zeigen, dass für die Messzeitpunkte t2, t3 und t4 stationarity vorliegt und auch die Faktor-Kovarianzen über die Zeit stabil bleiben.

Für Multigruppenanalyse 2 werden die gleichen hierarchisch abgestuften Modelle verwendet wie für Multigruppenanalyse 1. In Tabelle 34 sind die Modell-Fits für das Modell 1 für alle drei Datensätze aufgelistet.

Tabelle 34. Ergebnisse für Modell 1 (Test auf konfigurale Invarianz) der Multigruppenanalyse 2 für alle Datensätze

	Datensatz		
	L1	L2	L3
Modell 1	$\chi^2 = 35.51$, df = 32, $p = .31$, CFI = 1.00, RMSEA = .02, p-close = .99, $\chi^2/\text{df} = 1.11$	$\chi^2 = 62.10$, df = 32, $p < .01$, CFI = .98, RMSEA = .04, p-close = .97, $\chi^2/\text{df} = 1.94$	$\chi^2 = 75.40$, df = 32, $p < .001$, CFI = .98, RMSEA = .03, p-close = 1.00, $\chi^2/\text{df} = 2.36$

Anmerkung: *Modell 1.* ohne Restriktion, kein Parameter wurde zwischen den Messzeitpunkten gleichgesetzt

Die Ergebnisse zeigen, dass für alle Datensätze konfigurale Invarianz über alle vier Messzeitpunkte gegeben ist.

Auch für Multigruppenanalyse 2 werden die Ergebnisse aller weiteren Analysen nur für den Datensatz L1 dargestellt, die Ergebnisse unterscheiden sich nicht in den Datensätzen L1 und L3.

Ein genesteter Modellvergleich zeigt, dass sich Modell 2 nicht signifikant von Modell 1 unterscheidet ($\Delta\chi^2 = 7.48$, $df = 12$, $p = .82$). Somit lässt sich auch für die latenten Faktoren psychisches Wohlbefinden und körperliche Beschwerden über alle vier Messzeitpunkte zeigen, dass stationarity gegeben ist. Modell 3 zeigt dagegen eine schlechtere Datenanpassung als Modell 2 ($\Delta\chi^2 = 25.78$, $df = 9$, $p < .05$). Wird zugelassen, dass die Varianzen in beiden latenten Faktoren zu t1 frei variieren können, so ist dieses modifizierte Modell nicht mehr signifikant schlechter angepasst ($\Delta\chi^2 = 6.81$, $df = 12$, $p = .82$). Bei näherer Betrachtung der Varianzen zeigt sich, dass zu t1 die Varianz in beiden latenten Faktoren geringer ausfällt als zu allen späteren Messzeitpunkten.

Resümierend zeigen beide Multigruppenanalysen, dass stationarity für die Messmodelle gegeben ist, was für alle weiteren Analysen unbedingte Voraussetzung ist (Meredith & Horn, 2001; Vandenberg & Lance, 2000).

Die Überprüfung der Messmodelle wird für alle Messzeitpunkte und alle Datensätze mit der Bootstrap-Methode⁵⁰ erneut geschätzt. Da die Bootstrap-Methode in AMOS nicht in Kombination mit der FIML-Methode verwendet werden kann, wurden fehlende Werte in den Item-Parcels mit der EM-Methode geschätzt und ersetzt. In Tabelle 35 sind die mit der Bootstrap-Methode ermittelten Verzerrungen in den Parameterschätzern für alle Messzeitpunkte und Datensätze zusammengefasst.

⁵⁰ Es wurden für alle Analysen jeweils 250 Bootstrap-Stichproben verwendet (Nevitt & Hancock, 2001).

Tabelle 35. Ergebnisse der Berechnung der Messmodelle mit der Bootstrap-Methode für alle Messzeitpunkte und alle Datensätze

Messzeitpunkt	Datensatz											
	L1				L2				L3			
	t1	t2	t3	t4	t1	t2	t3	t4	t1	t2	t3	t4
Mittlere Zunahme der Standardfehler der (unstandardisierten) Faktorladungen	32.06%	43.88%	38.80%	10.05%	23.94%	28.92%	19.65%	13.81%	10%	20.58%	19.27%	16,42%
Mittlere Zunahme der Standardfehler der Faktorkovarianzen	22.31%	12.75%	22.31%	11.16	14.44%	9.24%	12.84%	6.01%	1.15%	1.28%	12.14%	12.83%
Mittlerer Bias in den Schätzern der (unstandardisierten) Faktorladungen	.03	.03	.04	.01	.02	.01	.01	.01	.01	.01	.01	.01
Mittlerer Bias in den Schätzern der Faktor-kovarianzen	.00	.01	.01	.01	.003	.00	.00	.00	.00	.00	.00	.00

Die Ergebnisse der Überprüfung der Messmodelle mit der Bootstrap-Methode zeigen, dass für alle Messzeitpunkte und alle Datensätze der Bias in den Parameterschätzern in einem akzeptablen Bereich liegt, der mittlere Bias liegt maximal bei $M = .04$. Darüber hinaus zeigen die bias-korrigierten Konfidenzintervalle, dass alle Faktorladungen größer als Null sind. Auch die bias-korrigierten Konfidenzintervalle für die mit der ML-Methode signifikanten Faktorinterkorrelationen schließen nicht die Null ein. Insgesamt kann resümiert werden, dass die Verzerrungen, die möglicherweise durch die Abweichung der Daten von der multivariaten Normalverteilung verursacht sein können, relativ gering ausfallen und auf die Parameterschätzer nur geringen Einfluss ausüben. Alle Parameter, die unter Verwendung der ML-Methode statistisch signifikant sind, sind dies auch bei Verwendung der Bootstrap-Methode.⁵¹

In einem abschließenden Schritt wurde mittels einer weiteren Multigruppenanalyse überprüft, ob es systematische Verzerrungen zwischen den drei Datensätzen gibt. Sollten Parameterschätzer zwischen den Datensätzen abweichen, so wäre dies ein Hinweis, dass durch den Dropout bzw. durch unterschiedliche Typen von Continuents die Ergebnisse beeinflusst sind (s.o.). Jeweils für jeden Messzeitpunkt wurden mehrere hierarchisch abgestufte Modelle berechnet, wobei ausgehend von einem Modell ohne Restriktion die Modelle immer restriktiver werden:

- *Modell 1.* ohne Restriktion (Alle Parameter können in allen Datensätzen variieren, Basismodell)
- *Modell 2.* Faktorladungen sind in allen Datensätzen gleich
- *Modell 3.* wie 2 und Messintercepts sind gleich
- *Modell 4.* wie 3 und Interkorrelationen sind gleich
- *Modell 5.* wie 4 und Fehlervarianzen sind gleich

In Tabelle 36 sind die Ergebnisse zusammengefasst.

⁵¹ Die stärkeren Verzerrungen in dem Datensatz L1 sind auf das geringe N zurückzuführen (Nevitt & Hancock, 2001).

Table 36. Ergebnisse der Multigruppenanalyse zur Überprüfung von Verzerrungseffekten durch unterschiedliche Typen von Continuern

Messzeitpunkt	Modell 1	Modell 2	Modell 3	Modell 4	Modell 5
	Fit-Indizes	χ^2 -Differenztest			
t1	$\chi^2 = 150.13$, $df = 72$, $p < .001$, CFI = .96, RMSEA = .04, p-close = .87, $\chi^2/df = 2.09$	$\Delta\chi^2 = 7.65$, $df = 12$, $p = .81$	$\Delta\chi^2 = 9.44$, $df = 18$, $p = .95$	$\Delta\chi^2 = 10.96$, $df = 12$, $p = .53$	$\Delta\chi^2 = 6.54$, $df = 18$, $p = .99$
t2	$\chi^2 = 749.77$, $df = 411$, $p < .001$, CFI = .93, RMSEA = .04, p-close = 1.00, $\chi^2/df = 1.82$	$\Delta\chi^2 = 11.20$, $df = 26$, $p = 1.00$	$\Delta\chi^2 = 12.84$, $df = 38$, $p = 1.00$	$\Delta\chi^2 = 14.48$, $df = 42$, $p = 1.00$	$\Delta\chi^2 = 23.60$, $df = 38$, $p = 1.00$
t3	$\chi^2 = 730.95$, $df = 411$, $p < .001$, CFI = .93, RMSEA = .04, p-close = 1.00, $\chi^2/df = 1.78$	$\Delta\chi^2 = 7.11$, $df = 26$, $p = 1.00$	$\Delta\chi^2 = 12.43$, $df = 38$, $p = 1.00$	$\Delta\chi^2 = 8.35$, $df = 42$, $p = 1.00$	$\Delta\chi^2 = 12.41$, $df = 38$, $p = 1.00$
t4	$\chi^2 = 652.14$, $df = 411$, $p < .001$, CFI = .94, RMSEA = .03, p-close = 1.00, $\chi^2/df = 1.59$	$\Delta\chi^2 = 8.67$, $df = 26$, $p = 1.00$	$\Delta\chi^2 = 5.65$, $df = 38$, $p = 1.00$	$\Delta\chi^2 = 7.63$, $df = 42$, $p = 1.00$	$\Delta\chi^2 = 9.76$, $df = 38$, $p = 1.00$

Anmerkungen: *Modell 1.* ohne Restriktion, kein Parameter wurde zwischen den Datensätzen gleichgesetzt; *Modell 2.* Faktorladungen sind in allen Datensätzen gleich; *Modell 3.* wie 2 und Messintercepts sind gleich; *Modell 4.* wie 3 und Interkorrelationen sind gleich; *Modell 5* wie 4 und Fehlervarianzen sind gleich

Die Ergebnisse zeigen, dass sich die Messmodelle zwischen den Datensätzen nicht unterscheiden, die Fit-Indizes für Modell 1 der Multigruppenanalyse liegen zu jedem Messzeitpunkt in einem befriedigenden Bereich, alle weiteren hierarchisch gestuften Modellvergleiche weisen auf keinerlei Abweichungen zwischen den Datensätzen hin. Darüber hinaus weist das restriktivste Modell 5 die besten Fit-Indizes auf (t1: $\chi^2 = 184.72$, $df = 132$, $p < .01$, CFI = .97, RMSEA = .03, p-close = 1.00, $\chi^2/df = 1.40$; t2: $\chi^2 = 811.89$, $df = 555$, $p < .001$, CFI = .95, RMSEA = .03, p-close = 1.00, $\chi^2/df = 1.46$; t3: $\chi^2 = 771.24$, $df = 555$, $p < .001$, CFI = .95, RMSEA = .02, p-close = 1.00, $\chi^2/df = 1.39$; t4: $\chi^2 = 683.84$, $df = 555$, $p < .001$, CFI = .97, RMSEA = .02, p-close = 1.00, $\chi^2/df = 1.23$). Somit deuten die Ergebnisse auf keinerlei Verzerrungen durch selektiven Dropout hin.

7.2.5 Überprüfung des Stressmodells im Querschnitt

In diesem Abschnitt geht es um die Überprüfung des in Studie 1 modifizierten Stressmodells (s. Abbildung 4). Eine Bestätigung des Modells zu den Messzeitpunkten 2, 3 und 4 in allen drei Datensätzen würde eine Kreuzvalidierung darstellen und die in Studie 1 vorgenommene Modifikation stützen. Tabelle 37 gibt die Ergebnisse der Strukturgleichungsanalysen zu den Messzeitpunkten 2, 3 und 4 wieder.

Tabelle 37. Ergebnisse der Überprüfung des Stressmodells zu den Messzeitpunkten 2, 3 und 4 unter Verwendung von SEM für alle drei Datensätze

Messzeitpunkt	Datensatz		
	L1	L2	L3
t2	$\chi^2 = 325.60$, df = 196, $p < .01$, CFI = .88, RMSEA = .08, p-close < .001, $\chi^2/df = 1.61$	$\chi^2 = 395.13$, df = 196, $p < .01$, CFI = .90, RMSEA = .07, p-close < .01, $\chi^2/df = 2.02$	$\chi^2 = 444.53$, df = 196, $p < .01$, CFI = .91, RMSEA = .07, p-close < .01, $\chi^2/df = 2.27$
t3	$\chi^2 = 311.66$, df = 196, $p < .01$, CFI = .89, RMSEA = .08, p-close < .01, $\chi^2/df = 1.59$	$\chi^2 = 403.56$, df = 196, $p < .01$, CFI = .90, RMSEA = .08, p-close < .01, $\chi^2/df = 2.06$	$\chi^2 = 384.25$, df = 196, $p < .01$, CFI = .92, RMSEA = .06, p-close = .09, $\chi^2/df = 1.96$
t4	$\chi^2 = 314.08$, df = 196, $p < .01$, CFI = .89, RMSEA = .08, p-close < .01, $\chi^2/df = 1.60$	$\chi^2 = 333.08$, df = 196, $p < .01$, CFI = .92, RMSEA = .06, p-close = .05, $\chi^2/df = 1.71$	$\chi^2 = 377.79$, df = 196, $p < .01$, CFI = .91, RMSEA = .06, p-close = .13, $\chi^2/df = 1.93$

Zunächst zeigen die Ergebnisse, dass die Fit-Indizes nur knapp befriedigend ausfallen. Gerade im Datensatz L1 liegt der CFI in einem unbefriedigenden Bereich, in den Datensätzen L2 und L3 liegen die Werte der Fit-Indizes etwas höher. Ein augenscheinlicher Vergleich der Strukturgewichte und Faktor-Kovarianzen mit Studie 1 zeigt nur wenige Abweichungen: Im Datensatz L1 wird jeweils zu t2, t3 und t4 das Strukturgewicht von Neurotizismus auf soziale Unterstützung nicht signifikant. Zusätzlich werden zu t3 das Strukturgewicht von Bedrohungsbewertung zu Coping und zu t4 die Strukturgewichte von Selbstwirksamkeitserwartungen auf Coping und von Neurotizismus auf psychisches Wohlbefinden nicht signifikant.⁵² Im Datensatz L2 wird das Strukturgewicht von Bewertung auf Coping zu t2 und das Strukturgewicht von Selbstwirksamkeitserwartungen auf Coping zu t3 nicht signifikant. Datensatz L3 enthält keine Abweichungen.

⁵² Einseitige Signifikanztestung.

Die Ergebnisse einer Multigruppenanalyse bestätigen die augenscheinlich nur geringen Abweichungen zwischen dem Längsschnitt- und Querschnittsdatsatz. Für alle drei Längsschnittdatensätze wurde das Querschnittsmodell zum zweiten Messzeitpunkt mit der Kalibrierungsstichprobe aus Studie 1 verglichen. In hierarchisch abgestuften Modellen wurden die Faktorladungen, Strukturgewichte und Interkorrelationen miteinander verglichen, die Ergebnisse der Analyse sind in Tabelle 38 dokumentiert:

- *Modell 1.* ohne Restriktion; kein Parameter wurde zwischen den Daten zu Messzeitpunkt 2 der drei längsschnittlichen Datensätzen und der Kalibrierungsstichprobe aus Studie 1 gleichgesetzt
- *Modell 2.* Faktorladungen sind gleich
- *Modell 3.* wie 2 und Strukturgewichte sind gleich
- *Modell 4.* wie 3 und Interkorrelationen sind gleich

Tabelle 38. Ergebnisse der Multigruppenanalyse zum Vergleich der drei Längsschnittdatensätze (L1, L2, L3) zum Messzeitpunkt 2 und der Kalibrierungsstichprobe aus Studie 1

Längsschnitt-Datensatz	Modell 1	Modell 2	Modell 3	Modell 4
	Fit-Indizes	χ^2 -Differenztest		
L1	$\chi^2 = 679.94$, $df = 392$, $p < .001$, CFI = .93, RMSEA = .04, p-close = .99, $\chi^2/df = 1.74$	$\Delta\chi^2 = 22.85$, $df = 15$, $p = .09$	$\Delta\chi^2 = 6.83$, $df = 11$, $p = .81$	$\Delta\chi^2 = 4.15$, $df = 2$, $p = .13$
L2	$\chi^2 = 758.49$, $df = 392$, $p < .001$, CFI = .93, RMSEA = .04, p-close = 1.00, $\chi^2/df = 1.94$	$\Delta\chi^2 = 20.44$, $df = 15$, $p = .16$	$\Delta\chi^2 = 19.38$, $df = 11$, $p = .06$	$\Delta\chi^2 = 5.59$, $df = 2$, $p = .06$
L3	$\chi^2 = 807.66$, $df = 392$, $p < .001$, CFI = .93, RMSEA = .04, p-close = 1.00, $\chi^2/df = 2.06$	$\Delta\chi^2 = 16.23$, $df = 15$, $p = .37$	$\Delta\chi^2 = 29.01$, $df = 11$, $p < .01$	$\Delta\chi^2 = 15.56$, $df = 2$, $p < .001$

Anmerkungen: *Modell 1.* ohne Restriktion, kein Parameter wurde zwischen den Datensätzen gleichgesetzt; *Modell 2.* Faktorladungen sind in allen Datensätzen gleich; *Modell 3.* wie 2 und Strukturgewichte sind gleich; *Modell 4.* wie 3 und Interkorrelationen sind gleich

Für die Datensätze L2 und L1 ergeben sich keine statistisch signifikanten Abweichungen von den Ergebnissen der Kalibrierungsstichprobe. Für den Datensatz L3 zeigt sich allerdings, dass es signifikante Unterschiede in den Strukturgewichten und Interkorrelationen gibt.

Zurückzuführen sind die Unterschiede auf einen höheren Zusammenhang zwischen Coping und den Ergebnisvariablen, psychischem Wohlbefinden und dem Ausmaß an körperlichen Beschwerden, in dem Datensatz L3, während der Effekt von Neurotizismus auf die beiden Ergebnisvariablen hier geringer ausfällt. Des Weiteren ist die Interkorrelation zwischen der Ereignis- und Ressourcenschätzung im Datensatz L3 höher als in der Kalibrierungsstichprobe aus Studie 1. Werden die entsprechenden Strukturgewichte ($\Delta\chi^2 = 7.60$, $df = 7$, $p = .37$) und Interkorrelationen ($\Delta\chi^2 = 0.79$, $df = 1$, $p = .38$) frei gesetzt, zeigen sich keine signifikanten Unterschiede mehr zwischen den beiden Datensätzen. Zusammenfassend zeigen sich bei einer querschnittlichen Betrachtung des Stressmodells nur geringe Abweichungen zu den Ergebnissen aus Studie 1. Da es keine Abweichungen in den Faktorladungen der Parcels zwischen der Kalibrierungsstichprobe aus Studie 1 und den drei Längsschnittdatensätzen gibt, kann davon ausgegangen werden, dass in beiden Studien die inhaltliche Bedeutung der Modellvariablen gleich ist.

Um die Stabilität des Modells über die Messzeitpunkte zu berechnen, wurde eine Multi-gruppenanalyse jeweils getrennt für die Datensätze berechnet. Dabei wurden wiederum mehrere hierarchisch abgestufte Modelle berechnet, wobei ausgehend von einem Modell ohne Restriktion die Modelle immer restriktiver werden:

- *Modell 1.* ohne Restriktion; kein Parameter wurde zwischen den Messzeitpunkten gleichgesetzt
- *Modell 2.* Faktorladungen zu allen Messzeitpunkten gleich
- *Modell 3.* wie 2 und Strukturgewichte sind gleich
- *Modell 4.* wie 3 und Interkorrelationen sind gleich
- *Modell 5.* wie 4 und Residualvarianz der Strukturvariablen ist gleich
- *Modell 6.* wie 5 und Fehlervarianzen sind gleich

Zunächst zeigt die Multigruppenanalyse für alle drei Datensätze, dass von einer Stabilität des Stressmodells ausgegangen werden kann (s. Tabelle 39).

Tabelle 39. Ergebnisse für Modell 1 der Multigruppenanalyse zum Test auf Stabilität des (querschnittlichen) Stressmodells für alle Datensätze

	Datensatz		
	L1	L2	L3
Modell 1	$\chi^2 = 941.33$, df = 588, $p < .001$, CFI = .89, RMSEA = .05, p-close = .86, $\chi^2/\text{df} = 1.60$	$\chi^2 = 1145.58$, df = 588, $p < .001$, CFI = .90, RMSEA = .04, p-close = 1.00, $\chi^2/\text{df} = 1.95$	$\chi^2 = 1206.56$, df = 588, $p < .001$, CFI = .91, RMSEA = .03, p-close = 1.00, $\chi^2/\text{df} = 2.05$

Anmerkung: *Modell 1.* ohne Restriktion, kein Parameter wurde zwischen beiden Stichproben gleichgesetzt

In allen drei Datensätzen zeigen alle weiteren hierarchisch abgestuften Modelle keine schlechtere Datenanpassung als Modell 1. Somit kann von der Stabilität des empirischen Stressmodells über die Messzeitpunkte t2, t3 und t4 ausgegangen werden.

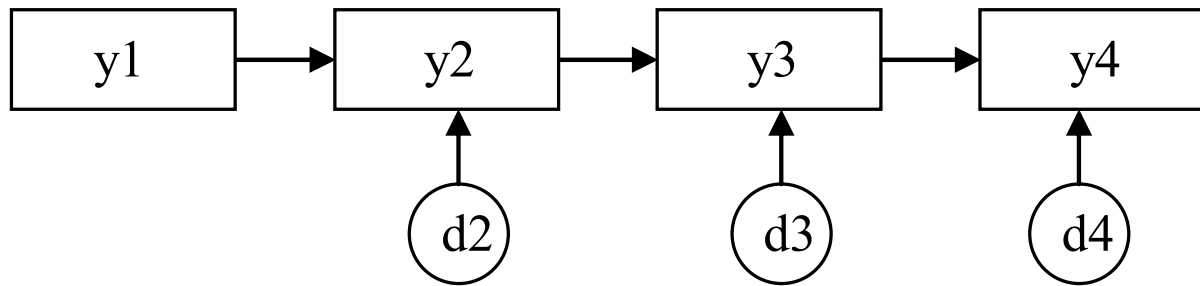
7.2.6 Überprüfung des Stressmodells im Längsschnitt

Wie oben beschrieben, wird zur Überprüfung des Stressmodells im Längsschnitt auf das so genannte ALT-Modell zurückgegriffen (Bollen & Curran, 2004; Curran & Bollen, 2001). Dabei wird für die Datenanalyse in Anlehnung an Curran und Bollen (2001) ein mehrschrittiges Vorgehen gewählt.

Zunächst wird in *Schritt 1* für alle Modellvariablen geprüft, ob sich die Mittelwerte über die Zeit verändern, wobei auf einen SEM-Ansatz zurückgegriffen wird: Für alle Modellvariablen werden jeweils zwei Modelle berechnet und miteinander verglichen. In Modell 1 können alle Mittelwerte und Kovarianzen zwischen den Wiederholungsmaßen variieren, in Modell 2 werden dagegen die Mittelwerte und Kovarianzen restringiert.

In *Schritt 2* wird ein ARCL-Modell (Simplex-Modell) für alle Modellvariablen berechnet. Der Unterschied zu Modell 1 ist hier, dass nun jeweils das vorherige Maß das zeitlich nächste Maß vorhersagt (s. Abbildung 8).

Abbildung 8. Beispiel für ein ARCL-Modell (Simplex-Modell) in Schritt 2 mit vier Wiederholungsmaßen



Auch hier werden ausgehend von einem Ausgangsmodell, in welchem die Strukturgewichte, der Mittelwert und die Varianz für den ersten Messzeitpunkt und die Intercepts und Residualvarianzen nicht restringiert sind, weitere Modelle berechnet, in denen Gleichheitsrestringtionen vorgenommen werden. Zunächst werden alle autoregressiven Parameter, anschließend die Residualvarianzen und abschließend die Intercepts restringiert.

Im *dritten Schritt* wird ein LGC-Modell mit einem Faktor berechnet, in dem nur der Intercept für das entsprechende Wiederholungsmaß geschätzt wird. Dieses Modell entspricht einer einfaktoriellen ANOVA mit Messwiederholung mit zufälligen Effekten (Bryk & Raudenbush, 1992). Der Unterschied zu dem Simplex-Modell liegt in der Annahme, dass alle Wiederholungsmaße durch einen den Maßen unterliegenden latenten Faktor beeinflusst sind und dass dieser Faktor für die beobachteten Kovarianzen und die Mittelwertsstruktur verantwortlich ist.

In *Schritt 4* wird das Modell aus Schritt 3 um einen weiteren Faktor ergänzt, der die mögliche systematische Veränderung über die Zeit in den Modellvariablen erklären soll. Dieses Modell entspricht dem weiter oben beschriebenen LGC-Modell (s. Abbildung 6).

In *Analyseschritt 5* wird für alle Modellvariablen getrennt ein ALT-Modell berechnet (s. Abbildung 7), welches das ARCL- und LGC-Modell kombiniert. Dieses Modell schätzt somit gleichzeitig zwei Komponenten von Veränderungen über die Zeit: Zum einen repräsentieren die latenten Faktoren die individuelle Variabilität in den kontinuierlich über die Zeit verlaufenden Veränderungen, zum anderen repräsentieren die autoregressiven Parameter Einflüsse auf Gruppenebene durch zeitlich vorher liegende Zeitpunkte (gleiche Effekte für alle Individuen).

Diese fünf Modelle werden wiederum miteinander verglichen und auf Basis der Ergebnisse wird für jede Modellvariable das jeweils am besten passende Modell verwendet. Für

Neurotizismus liegen nur Daten für den ersten Messzeitpunkt vor, somit wird für diese Modellvariable keiner der beschriebenen Analyseschritte vollzogen.

Im Anschluss an die Analysen wird abschließend das Stressmodell im Längsschnitt geprüft, in dem ein multivariates ALT-Modell berechnet wird, wobei sukzessive die einzelnen Modellvariablen in das Modell aufgenommen werden. Die genauen Analyseschritte werden an entsprechender Stelle beschrieben.

Aufgrund der geringen Stichprobengröße werden für die ARCL-Modelle nur die manifesten Variablen verwendet. Die Überprüfung der Messmodelle zu allen Messzeitpunkten und für alle Datensätze zeigt, dass die aus den einzelnen Indikatoren gebildeten latenten Faktoren stabil sind und der aus den jeweiligen Indikatoren gebildete Mittelwert interpretiert werden kann. Ebenso werden für alle LGC- und ALT-Modelle ebenfalls nur die Mittelwerte der einzelnen Variablen als Indikatoren verwendet. Das endgültige multivariate LGC-Modell wird aber auch mit latenten Indikatoren berechnet und mit den Ergebnissen, die mit manifesten Indikatoren ermittelt wurden, verglichen.

Wie weiter oben erwähnt, werden die beschriebenen Analyseschritte mit Datensatz L2 durchgeführt. Die Datensätze L1 und L3 dienen zur Absicherung der mit Datensatz L2 ermittelten Ergebnisse. In Kapitel 7.1.6.1 werden zunächst die Ergebnisse für Schritt 1 bis 5 beschrieben und miteinander verglichen, in Abschnitt 7.1.6.2 wird anschließend das multivariate LGC-Modell berechnet.

7.2.6.1 Berechnung der ARCL-, LGC- und ALT-Modelle für alle Modellvariablen

In Tabelle 40 sind Mittelwerte und Standardabweichungen für alle Modellvariablen zu allen Messzeitpunkten dargestellt.⁵³ Die Interkorrelationen zwischen den Variablen sind im Anhang für den Datensatz L2 dokumentiert (s. S. XXXVIII).

⁵³ Die Mittelwerte und Standardabweichungen wurden in SPSS berechnet und basieren auf, mit der EM-Methode ersetzten, fehlenden Werten in Datensatz L2.

Tabelle 40. Mittelwerte und Standardabweichungen für die Modellvariablen zu allen Messzeitpunkten (Datensatz L2)

Messzeitpunkt	t1		t2		t3		t4	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Neurotizismus	2.74	0.90	x	x	x	x	x	x
Ausmaß an sozialer Unterstützung	x	x	3.88	0.82	3.86	0.78	3.86	0.89
Ereigniseinschätzung	x	x	2.25	0.90	2.22	0.95	2.35	1.02
Selbstwirksamkeits-einschätzung	x	x	4.16	0.74	4.19	0.72	4.31	0.80
Coping	x	x	4.59	0.70	4.46	0.75	4.53	0.79
Körperliche Beschwerden	2.35	0.82	2.69	0.97	2.52	1.01	2.87	0.99
Psychisches Wohlbefinden	4.70	0.72	4.24	0.81	4.26	0.83	4.02	0.93

Anmerkung: x = Variable wurde zu diesem Messzeitpunkt nicht erhoben

Schritt 1

In einem ersten Schritt wurde für alle Modellvariablen die Stabilität der Mittelwerte geprüft, in dem jeweils zwei Strukturgleichungsmodelle berechnet wurden, die miteinander verglichen werden: Ein Modell ohne Restriktionen und ein Modell, in dem die Mittelwerte der jeweiligen Modellvariable zu den verschiedenen Messzeitpunkten gleich gesetzt wurden. Da das Modell ohne Restriktionen exakt identifiziert ist (d.h. die Anzahl der zu schätzenden Parameter entspricht der Anzahl der Daten) und somit keine Fit-Indizes angegeben werden können,

werden nur die Ergebnisse für die restringierten Modelle angegeben und bewertet. In Tabelle 41 sind die Ergebnisse für alle Modellvariablen aufgeführt.

Tabelle 41. Fit-Indizes für das restringierte Modell zur Überprüfung der Stabilität der Mittelwerte in allen Modellvariablen über die Zeit

Modellvariable	Fit-Indizes
Ausmaß an sozialer Unterstützung	$\chi^2 = 0.13$, $df = 2$, $p = .94$, CFI = 1.00, RMSEA = .00, p-close = .96, $\chi^2/df = 0.07$
Ereigniseinschätzung	$\chi^2 = 2.76$, $df = 2$, $p = .25$, CFI = 1.00, RMSEA = .05, p-close = .40, $\chi^2/df = 1.38$
Selbstwirksamkeitserwartungen	$\chi^2 = 2.16$, $df = 2$, $p = .34$, CFI = 1.00, RMSEA = .02, p-close = .49, $\chi^2/df = 1.08$
Coping	$\chi^2 = 8.88$, $df = 2$, $p < .05$, CFI = 0.97, RMSEA = .14, p-close < .05, $\chi^2/df = 4.44$
Körperliche Beschwerden	$\chi^2 = 55.08$, $df = 3$, $p < .001$, CFI = .84, RMSEA = .30, p-close < .001, $\chi^2/df = 18.36$
Psychisches Wohlbefinden	$\chi^2 = 80.27$, $df = 3$, $p < .001$, CFI = .60, RMSEA = .37, p-close < .001, $\chi^2/df = 26.76$

Die Ergebnisse zeigen, dass die Mittelwerte der Modellvariablen Ausmaß an sozialer Unterstützung, Ereigniseinschätzung und Selbstwirksamkeitserwartungen über die Zeit hinweg gleich sind, alle Fit-Indizes liegen in einem sehr guten Bereich. Dagegen sprechen die Fit-Indizes in den Variablen körperliche Beschwerden und psychisches Wohlbefinden gegen die Annahme, dass die Mittelwerte stabil bleiben, was auch mit Tabelle 40 korrespondiert. Wenn auch der CFI für das Mittelwertsmodell der Modellvariable Coping in einem guten Bereich liegt, spricht der RMSEA gegen die Annahme, dass das restringierte Modell auf die Daten passt. In Schritt 4 werden LGC-Modelle verwendet, um das Mittelwertsmodell in Abhängigkeit der Zeit zu analysieren.

Schritt 2

Zunächst werden aber in Schritt 2 für alle Modellvariablen ARCL- bzw. Simplex-Modelle berechnet. In Tabelle 42 sind die Ergebnisse zusammengefasst.

Tabelle 42. Fit-Indizes für das ARCL-Modell jeweils getrennt für die Modellvariablen (Modelle ohne Restriktionen)

Modellvariable	Fit-Indizes
Ausmaß an sozialer Unterstützung	$\chi^2 = 4.98$, $df = 1$, $p < .05$, CFI = .98, RMSEA = .15, p-close = .06, $\chi^2/df = 4.98$
Ereigniseinschätzung	$\chi^2 = 5.23$, $df = 1$, $p < .05$, CFI = .98, RMSEA = .15, p-close = .06, $\chi^2/df = 5.23$
Selbstwirksamkeitserwartungen	$\chi^2 = 28.83$, $df = 1$, $p < .001$, CFI = .88, RMSEA = .38, p-close < .001, $\chi^2/df = 28.83$
Coping	$\chi^2 = 6.74$, $df = 1$, $p < .01$, CFI = .97, RMSEA = .17, p-close < .05, $\chi^2/df = 6.74$
Körperliche Beschwerden	$\chi^2 = 46.51$, $df = 3$, $p < .001$, CFI = .87, RMSEA = .28, p-close < .001, $\chi^2/df = 15.50$
Psychisches Wohlbefinden	$\chi^2 = 13.37$, $df = 3$, $p < .01$, CFI = .95, RMSEA = .14, p-close < .05, $\chi^2/df = 4.46$

Den Ergebnissen aus Schritt 1 entsprechend, liegen die Fit-Indizes des ARCL-Modells für die Modellvariablen Coping, körperliche Beschwerden und psychisches Wohlbefinden in keinem akzeptablen Bereich. Aber auch für die übrigen Variablen sind die Fit-Werte nicht befriedigend: Der RMSEA liegt jeweils in einem unakzeptablen Bereich. Die Einführung von Restriktionen führt bei allen Modellvariablen entweder zu einer signifikanten Verschlechterung der Modellanpassung oder zu keiner Änderung in den Anpassungswerten. Die Ergebnisse deuten darauf hin, dass ARCL-Modelle die beobachteten Kovarianzen und Mittelwertstrukturen in den Modellvariablen nicht abbilden.

Schritt 3

In Schritt 3 soll überprüft werden, ob dem Mittelwertsmuster ein Intercept-Faktor zu Grunde liegt. Geprüft wird hiermit, ob den Wiederholungsmaßen eine stabile, zeitunabhängige Komponente gemein ist. In Tabelle 43 sind die Ergebnisse zusammengefasst.

Tabelle 43. Fit-Indizes für das einfaktorielle LGC-Modell jeweils getrennt für die Modellvariablen

Modellvariable	Fit-Indizes
Ausmaß an sozialer Unterstützung	$\chi^2 = 2.26$, $df = 4$, $p = .69$, CFI = 1.00, RMSEA = .00, p-close = .84, $\chi^2/df = 0.57$
Ereigniseinschätzung	$\chi^2 = 6.85$, $df = 4$, $p = .14$, CFI = .99, RMSEA = .06, p-close = .33, $\chi^2/df = 1.71$
Selbstwirksamkeitserwartungen	$\chi^2 = 5.05$, $df = 4$, $p = .28$, CFI = 1.00, RMSEA = .04, p-close = .50, $\chi^2/df = 1.26$
Coping	$\chi^2 = 15.01$, $df = 4$, $p < .01$, CFI = .95, RMSEA = .12, p-close < .05, $\chi^2/df = 3.75$
Körperliche Beschwerden	$\chi^2 = 83.91$, $df = 8$, $p < .001$, CFI = .77, RMSEA = .22, p-close < .001, $\chi^2/df = 10.49$
Psychisches Wohlbefinden	$\chi^2 = 115.59$, $df = 8$, $p < .001$, CFI = .44, RMSEA = .27, p-close < .001, $\chi^2/df = 14.45$

Für die Modellvariablen Ausmaß an sozialer Unterstützung, Ereigniseinschätzung und Selbstwirksamkeitserwartungen sind die Fit-Indizes in einem sehr guten Bereich. Diesen Variablen scheint eine zeitstabile Komponente zu unterliegen. Dagegen werden die Variablen Coping, körperliche Beschwerden und psychisches Wohlbefinden durch ein einfaktorielles LGC-Modell nur schlecht repräsentiert, die Fit-Indizes deuten auf eine große Modellabweichung hin.

In allen Modellvariablen ist der latente Intercept-Faktor durch einen signifikanten Mittelwert und eine signifikante Varianz gekennzeichnet, was auf eine den Variablen unterliegende stabile Komponente hindeutet.

Schritt 4

In Schritt 4 wird nun für jede Modellvariable ein zweifaktorielles LGC-Modell berechnet. Neben einem latenten Intercept-Faktor wird nun noch ein latenter Slope-Faktor hinzugefügt, der mögliche Änderungen über die Zeit abbilden soll. Standardmäßig wird zunächst für alle Variablen ein linearer Slope-Faktor angenommen, es werden aber auch je nach Mittelwertverlauf andere Veränderungsverläufe analysiert.⁵⁴ In Tabelle 44 sind die Ergebnisse zusammengefasst.

Tabelle 44. Fit-Indizes für das zweifaktorielle LGC-Modell jeweils getrennt für die Modellvariablen

Modellvariable	Fit-Indizes
Ausmaß an sozialer Unterstützung ¹	$\chi^2 = 0.12$, $df = 1$, $p = .73$, CFI = 1.00, RMSEA = .00, p-close = .78, $\chi^2/df = 0.12$
Ereigniseinschätzung ¹	$\chi^2 = 1.74$, $df = 1$, $p = .19$, CFI = 1.00, RMSEA = .06, p-close = .29, $\chi^2/df = 1.74$
Selbstwirksamkeitserwartungen ¹	$\chi^2 = 0.56$, $df = 1$, $p = .45$, CFI = 1.00, RMSEA = .00, p-close = .55, $\chi^2/df = 0.56$
Coping ²	$\chi^2 = 1.15$, $df = 2$, $p = .56$, CFI = 1.00, RMSEA = .00, p-close = .69, $\chi^2/df = 0.57$
Körperliche Beschwerden ³	$\chi^2 = 7.84$, $df = 4$, $p = .10$, CFI = .99, RMSEA = .07, p-close = .26, $\chi^2/df = 1.96$
Psychisches Wohlbefinden ⁴	$\chi^2 = 7.61$, $df = 4$, $p = .11$, CFI = .98, RMSEA = .07, p-close = .27, $\chi^2/df = 1.90$

Anmerkungen: 1 Linearer Slope; 2 Kurvilinearer Slope: $t_1 = 0$, $t_2 = 1$, $t_3 = 1$, $t_4 = 0.5$; 3 Kurvilinearer Verlauf: $t_1 = 0$, $t_2 =$ nicht restringiert, $t_3 =$ nicht restringiert, $t_4 = 1$; 4 Kurvilinearer Verlauf: $t_1 = 0$, $t_2 = 1$, $t_3 =$ nicht restringiert, $t_4 =$ nicht restringiert

Für alle Modellvariablen zeigt ein zweifaktorielles LGC-Modell eine gute Anpassung an die Daten. Allerdings wird in den Variablen Ausmaß an sozialer Unterstützung, Ereigniseinschätzung und Selbstwirksamkeitserwartungen weder der Mittelwert noch die Varianz signi-

⁵⁴ Ein linearer Slope-Faktor wird erzeugt, indem die Faktorladung einer Variablen zum ersten Messzeitpunkt auf null restringiert wird, zum zweiten Messzeitpunkt auf eins, zum dritten auf zwei und zum vierten auf drei. Andere Veränderungsverläufe können analysiert werden, indem die Faktorladungen auf andere Art und Weise restringiert werden (z.B. exponentieller Veränderungsverlauf: $t_1 = 0$; $t_2 = 1$, $t_3 = 2$, $t_4 = 4$).

fikant. Der Vergleich mit dem einfaktoriellen LGC-Modell zeigt für Ausmaß an sozialer Unterstützung ($\Delta\chi^2 = 2.14$, $df = 3$, $p > .25$), Ereigniseinschätzung ($\Delta\chi^2 = 5.11$, $df = 3$, $p > .10$) und Selbstwirksamkeitserwartungen ($\Delta\chi^2 = 4.49$, $df = 3$, $p = .20$) keine bedeutsame Verbesserung in der Datenanpassung (s. Tabelle 42). Dies zeigt, dass diesen Variablen kein Slope-Faktor unterliegt. Dagegen werden in den Variablen Coping, körperliche Beschwerden und psychisches Wohlbefinden⁵⁵ der Mittelwert und die Varianz signifikant. Im Vergleich zu dem einfaktoriellen LGC-Modell ist das zweifaktorielle Modell deutlich besser an die Daten angepasst (Coping: $\Delta\chi^2 = 13.87$, $df = 2$, $p < .01$; körperliche Beschwerden: $\Delta\chi^2 = 76.07$, $df = 4$, $p < .01$; psychisches Wohlbefinden: $\Delta\chi^2 = 107.98$, $df = 4$, $p < .01$; s. Tabelle 42). In diesen Variablen scheint es zu systematischen Veränderungen im Erhebungszeitraum gekommen zu sein (vgl. Mittelwerte für die Variablen in Tabelle 40).

Schritt 5

In einem letzten Schritt wird nun für jede Modellvariable eine Kombination aus ARCL- und LGC-Modell, ein so genanntes ALT-Modell überprüft. Dazu wird für die Modellvariablen Ausmaß an sozialer Unterstützung, Ereigniseinschätzung und Selbstwirksamkeitserwartungen dem einfaktoriellen LGC-Modell und für die Variablen Coping, körperliche Beschwerden und psychisches Wohlbefinden jeweils autoregressive Parameter hinzugefügt. In Tabelle 45 sind die Ergebnisse für die ALT-Modelle aufgeführt.

⁵⁵ Die Varianz wird allerdings für die Variable psychisches Wohlbefinden bei einseitiger Testung nur auf dem 10%-Niveau signifikant. Dies deutet auf eine geringe Variabilität in der Veränderung des psychischen Wohlbefindens hin.

Tabelle 45. Fit-Indizes für das ALT-Modell jeweils getrennt für die Modellvariablen

Modellvariable	Fit-Indizes
Ausmaß an sozialer Unterstützung	$\chi^2 = 2.19$, $df = 2$, $p = .34$, CFI = 1.00, RMSEA = .02, p-close = .49, $\chi^2/df = 1.09$
Ereigniseinschätzung	$\chi^2 = 3.38$, $df = 2$, $p = .18$, CFI = .99, RMSEA = .06, p-close = .33, $\chi^2/df = 1.69$
Selbstwirksamkeitserwartungen	$\chi^2 = 2.92$, $df = 2$, $p = .23$, CFI = 1.00, RMSEA = .05, p-close = .38, $\chi^2/df = 1.46$
Coping	$\chi^2 = 0.09$, $df = 1$, $p = .76$, CFI = 1.00, RMSEA = .00, p-close = .81, $\chi^2/df = 0.09$
Körperliche Beschwerden	$\chi^2 = 3.36$, $df = 1$, $p = .07$, CFI = .99, RMSEA = .11, p-close = .13, $\chi^2/df = 3.36$
Psychisches Wohlbefinden ¹	$\chi^2 = 0.94$, $df = 1$, $p = .33$, CFI = 1.00, RMSEA = .00, p-close = .44, $\chi^2/df = 0.94$

Anmerkung: 1 Die Varianz des Slope-Faktors wurde auf den aus der Analyse des zweifaktoriellen LGC-Modells ermittelten Wert restringiert, um einen Freiheitsgrad zu gewinnen.

Auch das ALT-Modell zeigt in allen Modellvariablen eine gute Anpassung an die Daten. Allerdings werden in allen Fällen die autorregressiven Parameter nicht signifikant, somit unterliegen allen Modellvariablen keine autoregressiven Parameter.

Vergleich der Modelle

Abschließend sollen die in den Schritten 2 bis 5 analysierten Modelle miteinander verglichen werden. Basierend auf diesem Vergleich werden die optimalen Modelle für jede Modellvariable ausgewählt und in die anschließende Berechnung des multivariaten ALT-Modells aufgenommen. In Tabelle 46 sind die Fit-Indizes für alle Modelle sowie die jeweiligen Modellvergleiche jeweils für alle Modellvariablen zusammengefasst.

Tabelle 46. Vergleich von ARCL-, ein- und zweifaktoriellem LGC-Modell und ALT-Modell für alle Modellvariablen

Skala	Schritt 2: ARCL-Modell	Schritt 3: einfaktorielles LGC-Modell ¹	Schritt 4: zweifaktorielles LGC-Modell ²	Schritt 5: ALT-Modell ³	Bestes Modell
Soziale Unterstützung	$\chi^2 = 4.98$, $df = 1$, $p < .05$, CFI = .98, RMSEA = .15, p-close = .06, $\chi^2/df = 4.98$	$\chi^2 = 2.26$; $df = 4$; $p = .69$; CFI = 1.00; RMSEA = .00; p-close = .84; $\chi^2/df = 0.57$ $\Delta\chi^2 = 2.72$, $df = 3$, $p > .25$	$\chi^2 = 0.12$; $df = 1$; $p = .73$; CFI = 1.00; RMSEA = .00; p-close = .78; $\chi^2/df = 0.12$ $\Delta\chi^2 = 2.14$, $df = 3$, $p > .25$	$\chi^2 = 2.19$; $df = 2$; $p < .34$; CFI = 1.00; RMSEA = .02; p-close = .49; $\chi^2/df = 1.09$ $\Delta\chi^2 = 2.07$., $df = 1$, $p > .25$	Einfaktorielles LGC-Modell
Ereigniseinschätzung	$\chi^2 = 5.23$, $df = 1$, $p < .05$, CFI = .98, RMSEA = .15, p-close = .06, $\chi^2/df = 5.23$	$\chi^2 = 6.85$; $df = 4$; $p = .14$; CFI = .99; RMSEA = .06; p-close = .33; $\chi^2/df = 1.71$ $\Delta\chi^2 = 1.63$, $df = 3$, $p > .25$	$\chi^2 = 1.74$; $df = 1$; $p = .19$; CFI = 1.00; RMSEA = .06; p-close = .29; $\chi^2/df = 1.74$ $\Delta\chi^2 = 5.11$, $df = 3$, $p > .10$	$\chi^2 = 3.38$; $df = 2$; $p < .18$; CFI = .99; RMSEA = .06; p-close = .33; $\chi^2/df = 1.69$ $\Delta\chi^2 = 1.64$, $df = 1$, $p < .10$	Einfaktorielles LGC-Modell
Selbstwirksamkeitserwartung	$\chi^2 = 28.83$, $df = 1$, $p < .001$, CFI = .88, RMSEA = .38, p-close < .001, $\chi^2/df = 28.83$	$\chi^2 = 5.05$; $df = 4$; $p = .28$; CFI = 1.00; RMSEA = .04; p-close = .50; $\chi^2/df = 1.26$ $\Delta\chi^2 = 23.78$, $df = 3$, $p < .01$	$\chi^2 = .56$; $df = 1$; $p = .45$; CFI = 1.00; RMSEA = .00; p-close = .55; $\chi^2/df = 0.56$ $\Delta\chi^2 = 4.49$, $df = 3$, $p > .10$	$\chi^2 = 2.92$; $df = 2$; $p < .23$; CFI = 1.00; RMSEA = .05; p-close = .38; $\chi^2/df = 1.46$, $\Delta\chi^2 = 2.36$., $df = 1$, $p > .25$	Einfaktorielles LGC-Modell
Coping	$\chi^2 = 6.74$, $df = 1$, $p < .01$, CFI = .97, RMSEA = .17, p-close < .05, $\chi^2/df = 6.74$	$\chi^2 = 15.01$; $df = 4$; $p < .01$; CFI = .95; RMSEA = .12; p-close = .03; $\chi^2/df = 3.75$ $\Delta\chi^2 = 8.27$, $df = 3$, $p < .01$	$\chi^2 = 1.15$; $df = 2$; $p = .56$; CFI = 1.00; RMSEA = .00; p-close = .69; $\chi^2/df = 0.57$ $\Delta\chi^2 = 13.87$, $df = 2$, $p < .01$	$\chi^2 = 0.09$, $df = 1$, $p = .76$, CFI = 1.00, RMSEA = .00, p-close = .81, $\chi^2/df = 0.09$ $\Delta\chi^2 = 1.06$., $df = 1$, $p > .25$	Zweifaktorielles LGC-Modell

Anmerkung: 1 Vergleich mit Modell aus Schritt 2; 2 Vergleich mit Modell aus Schritt 3; 3 Vergleich mit Modell aus Schritt 4

Fortsetzung der Tabelle auf der nächsten Seite

Fortsetzung Tabelle 46

Skala	Schritt 2: ARCL-Modell	Schritt 3: einfaktorielles LGC-Modell ¹	Schritt 4: zweifaktorielles LGC-Modell ²	Schritt 5: ALT-Modell ³	Bestes Modell
Körperliche Beschwerden	$\chi^2 = 13.37$, df = 3, $p < .01$, CFI = .95, RMSEA = .14, p-close < .05, $\chi^2/df = 4.46$	$\chi^2 = 83.91$; df = 8; $p < .01$; CFI = .77; RMSEA = .22; p-close < .001; $\chi^2/df = 10.49$ $\Delta\chi^2 = 70.54$, df = 5, $p < .01$	$\chi^2 = 7.84$; df = 4; $p < .10$; CFI = .99; RMSEA = .07; p-close = .26; $\chi^2/df = 1.96$ $\Delta\chi^2 = 76.07$, df = 4, $p < .01$	$\chi^2 = 3.36$; df = 1; $p < .07$; CFI = .99; RMSEA = .11; p-close = .13; $\chi^2/df = 3.36$ $\Delta\chi^2 = 4.48$, df = 3, $p > .10$	Zweifaktorielles LGC-Modell
Psychisches Wohlbefinden	$\chi^2 = 46.51$, df = 3, $p < .001$, CFI = .87, RMSEA = .28, p-close < .001, $\chi^2/df = 15.50$	$\chi^2 = 115.59$; df = 8; $p < .01$; CFI = .44; RMSEA = .27; p-close < .001; $\chi^2/df = 14.45$ $\Delta\chi^2 = 69.08$, df = 5, $p < .01$	$\chi^2 = 7.61$; df = 4; $p < .11$; CFI = .98; RMSEA = .07; p-close = .27; $\chi^2/df = 1.90$ $\Delta\chi^2 = 107.98$, df = 4, $p < .01$	$\chi^2 = 0.94$, df = 1, $p = .33$, CFI = 1.00, RMSEA = .00, p-close = .44, $\chi^2/df = 0.94$ $\Delta\chi^2 = 6.67$, df = 3, $p < .10$	Zweifaktorielles LGC-Modell

Anmerkungen: 1 Vergleich mit Modell aus Schritt 2; 2 Vergleich mit Modell aus Schritt 3; 3 Vergleich mit Modell aus Schritt 4

Für die Modellvariablen Ausmaß an sozialer Unterstützung, Ereigniseinschätzung und Selbstwirksamkeitserwartungen zeigt sich, dass jeweils ein einfaktorielles LGC-Modell als an die Daten am besten angepasst ist. Zwar ergibt der Vergleich mit einem ARCL-Modell mit Ausnahme für das Selbstwirksamkeitsmaß keinen signifikanten Unterschied (bzw. Verbesserung), allerdings ist der RMSEA für das ARCL-Modell der Variablen Ausmaß an sozialer Unterstützung und Ereigniseinschätzung inakzeptabel ($RMSEA > .10$; vgl. Schermelleh-Engel et al., 2003). Auch im Vergleich zu einem zweifaktoriellen LGC-Modell und einem ALT-Modell zeigt sich das einfaktorielle LGC-Modell für diese Variablen nach dem χ^2 -Differenztest nicht überlegen. Allerdings werden jeweils für alle Variablen weder im zweifaktoriellen LGC-Modell der Mittelwert und die Varianz des Slope-Faktors signifikant noch in einem ALT-Modell die autoregressiven Parameter. Somit kann für diese drei Modellvariablen festgehalten werden, dass ihnen jeweils eine gemeinsame Komponente zu Grunde liegt und über den Erhebungszeitraum sich keine bedeutsamen Veränderungen, weder inter- noch intraindividuell ergeben.

Für die weiteren Modellvariablen, also Coping, körperliche Beschwerden und psychisches Wohlbefinden bewährt sich hingegen ein zweifaktorielles LGC-Modell. Ein zweifaktorielles LGC-Modell ist jeweils dem einfaktoriellen LGC-Modell überlegen. Zwar zeigt ein ALT-Modell für alle drei Variablen keine schlechtere Datenanpassung, allerdings werden die autoregressiven Parameter nicht signifikant. Diesen drei Modellvariablen unterliegt also zu den Messzeitpunkten eine gemeinsame Komponente, gleichzeitig lässt sich über den Erhebungszeitraum eine bedeutsame Veränderung feststellen, die zusätzlich für die Befragten unterschiedlich stark ausfällt. Einschränkend muss hier allerdings festgehalten werden, dass die Varianz des Slope-Faktors für die Variable psychisches Wohlbefinden nur marginal signifikant ist.

Im Anhang sind für alle Modellvariablen alle Parameterschätzer des jeweils gewählten Modells dokumentiert (s. S. XLf; Tabellen A22 und A23)

7.2.6.2 Das multivariate ALT-Modell

Die Überprüfung des Stressmodells erfolgt wie oben begründet mit einem multivariaten ALT-Modell. Zwar werden keine autoregressiven Parameter mehr aufgenommen, da sich im vorangegangenen Kapitel gezeigt hat, dass sich für keine der Skalen solche Beziehungen empirisch zeigen. Aber es werden cross-lagged-Beziehungen parametrisiert, d.h. es wird überprüft, ob eine Modellvariable zu einem gegebenen Zeitpunkt durch eine andere Modell-

variable zu einem vorherigen Zeitpunkt vorhergesagt wird. Auch für diese Analyse ist das Vorgehen mehrschrittig, indem die Modellvariablen sukzessive in das ALT-Modell aufgenommen werden.⁵⁶ In folgender Reihenfolge werden die einzelnen Modellvariablen in das Gesamtmodell aufgenommen:

Schritt 1. Bivariates ALT-Modell mit körperlichen Beschwerden (zweifaktorielles LGC-Modell) und psychischem Wohlbefinden (zweifaktorielles LGC-Modell)

Schritt 2. Neurotizismus (Skala Messzeitpunkt 1)

Schritt 3. Coping (zweifaktorielles LGC-Modell)

Schritt 4. Ereigniseinschätzung (einfaktorielles LGC-Modell) und Selbstwirksamkeitserwartungen (einfaktorielles LGC-Modell)

Schritt 5. Ausmaß an sozialer Unterstützung (einfaktorielles LGC-Modell)

Innerhalb jeden Analyseschrittes wird das Modell nach folgender Regel optimiert, bevor jeweils der nächste Analyseschritt folgt: Zunächst werden (1) alle nicht signifikanten cross-lagged-Parameter sukzessive aus dem Modell herausgenommen, wobei jeweils überprüft wird, ob sich das Modell hierdurch signifikant verschlechtert. Anschließend werden (2) alle nicht signifikanten Interkorrelationen zwischen den Maßen der Modellvariablen zu einem Messzeitpunkt (also z.B. der Korrelation zwischen psychischem Wohlbefinden zu t3 und körperlichen Beschwerden zu t3) sukzessive aus dem Modell entfernt. In einem letzten Schritt werden alle nicht signifikanten Pfade zwischen den Intercept- und/oder Slope-Faktoren entfernt. In jedem Analyseschritt werden zunächst alle möglichen cross-lagged-Parameter, Interkorrelation und Pfade von der neu aufgenommenen Modellvariablen zugelassen, wobei im Anschluss daran wiederum sukzessive, in der beschriebenen Reihenfolge, nicht signifikante Parameter aus dem Modell herausgenommen werden.

Das durch diese verschiedenen Analyseschritte endgültig ermittelte längsschnittliche Stressmodell wird abschließend auch mit latenten Indikatoren für die jeweiligen Modellvariablen zu einem Messzeitpunkt geschätzt.

⁵⁶ Ein alternatives Vorgehen, in dem alle Modellvariablen zugleich aufgenommen werden, führt zu einem gleichen Endresultat.

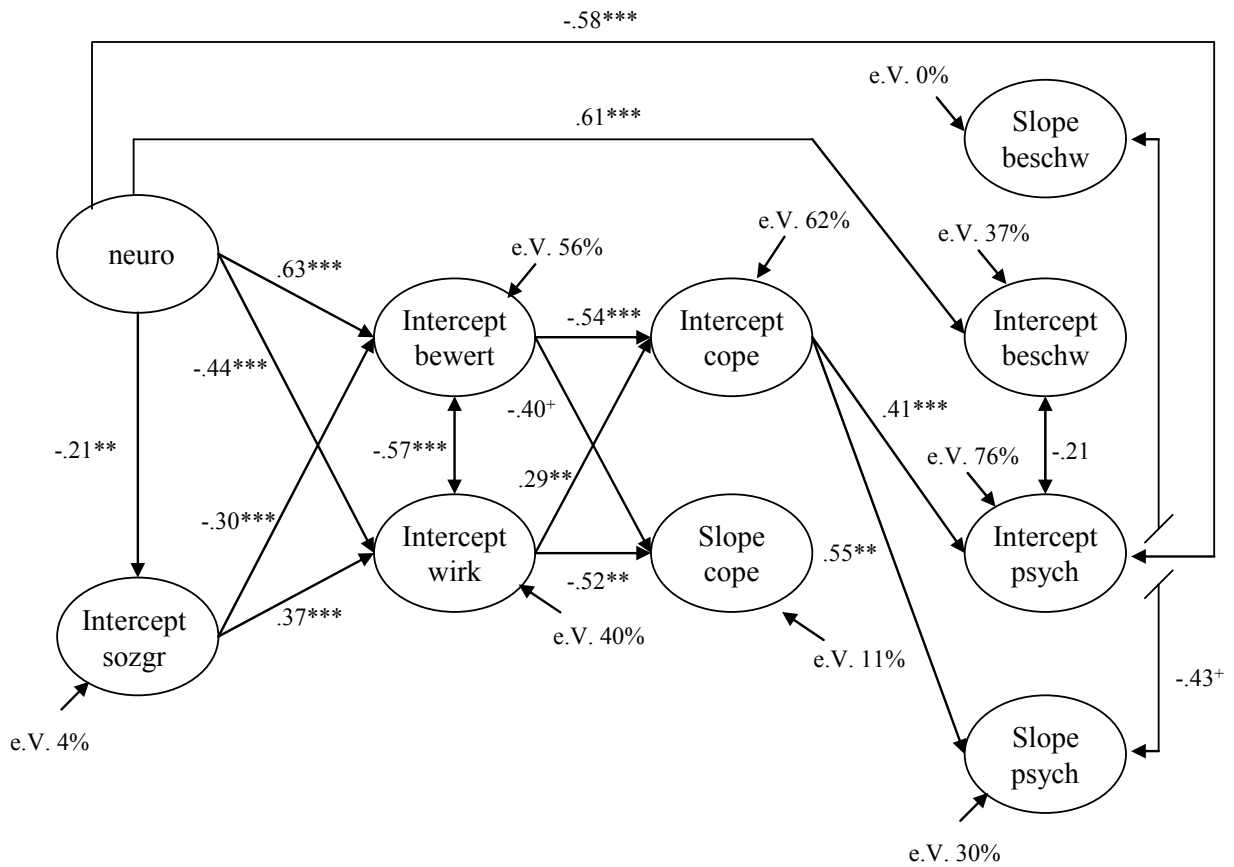
Tabelle 47 fasst die Fit-Indizes nach jedem der oben beschriebenen Analyseschritte 1 bis 5 zusammen.

Tabelle 47. Fit-Indizes für die zu jedem Analyseschritt optimierten ALT-Modelle (Datensatz L2)

Analyseschritt	Fit-Indizes
1 (körperliche Beschwerden und psychisches Wohlbefinden)	$\chi^2 = 28.05$, $df = 19$, $p = .08$, CFI = .99, RMSEA = .05, p-close = .46, $\chi^2/df = 1.48$
2 (Neurotizismus)	$\chi^2 = 31.58$, $df = 27$, $p = .25$, CFI = .99, RMSEA = .03, p-close = .78, $\chi^2/df = 1.17$
3 (Coping)	$\chi^2 = 57.82$, $df = 48$, $p = .16$, CFI = .99, RMSEA = .03, p-close = .83, $\chi^2/df = 1.20$
4 (Ereigniseinschätzung und Selbstwirksamkeitserwartungen)	$\chi^2 = 206.69$, $df = 123$, $p < .001$, CFI = .96, RMSEA = .06, p-close = .12, $\chi^2/df = 1.68$
5 (Ausmaß an sozialer Unterstützung)	$\chi^2 = 268.13$, $df = 173$, $p < .001$, CFI = .96, RMSEA = .05, p-close = .30, $\chi^2/df = 1.55$

Die Ergebnisse zeigen, dass das Modell nach jedem Analyseschritt gut an die Daten angepasst ist, es sind keinerlei Probleme bei der Parameterschätzung aufgetreten. In Abbildung 9 ist das Endmodell schematisch dargestellt, es sind nur die strukturellen Beziehungen zwischen den latenten Modellvariablen wiedergegeben. In Tabelle 48 sind alle die Intercepts und Varianzen der latenten Intercept- und Slope-Faktoren sowie der Skala Neurotizismus des multivariaten ALT-Modells zusammengefasst.

Abbildung 9. Schematische Darstellung der strukturellen Beziehungen zwischen den Modellvariablen für das multivariate ALT-Modell (standardisierte Parameter)



Anmerkungen: + $p < .10$; ** $p < .01$; *** $p < .001$; e.V. = erklärte Varianz; neuro = Neurotizismus; sozgr = Soziale Unterstützung; bewert = Ereignisseinschätzung; wirk = Selbstwirksamkeitserwartungen; cope = Coping; beschw = körperliche Beschwerden; psych = psychisches Wohlbefinden

Neurotizismus wurde nur zum ersten Messzeitpunkt erfasst, der latente Neurotizismusfaktor gibt somit den mittleren Neurotizismuswert für die Stichprobe sowie die Varianz in der Neurotizismusausprägung wieder. Die mit Intercept bezeichneten latenten Faktoren stellen die gemittelte Ausprägung in den entsprechenden Modellvariablen zu Beginn der Befragung dar. Die mit Slope bezeichneten latenten Faktoren zeigen dagegen die Veränderungen in den entsprechenden Modellkomponenten über die Zeit an.

Tabelle 48. Intercepts und Varianzen der latenten Intercept- und Slope-Faktoren sowie der Skala Neurotizismus des multivariaten ALT-Modells (Datensatz L2)

	Intercepts		
	Schätzer	Standardfehler	p
Neurotizismus	2.74	.06	< .001
Intercept soziale Unterstützung	4.28	.17	< .001
Intercept Ereignisseinschätzung	2.13	.36	< .001
Intercept Selbstwirksamkeitserwartungen	3.64	.34	< .001
Intercept Coping	4.46	.62	< .001
Slope Coping	1.34	.70	= .055
Intercept körperliche Beschwerden	1.03	.16	< .001
Slope körperliche Beschwerden	0.49	.06	< .001
Intercept psychisches Wohlbefinden	4.05	.41	< .001
Slope psychisches Wohlbefinden	-1.27	.33	< .001
	Varianz		
	Schätzer	Standardfehler	p
Neurotizismus	.76	.08	< .001
Intercept soziale Unterstützung	.39	.05	< .001
Intercept Ereignisseinschätzung	.26	.03	< .001
Intercept Selbstwirksamkeitserwartungen	.26	.04	< .001
Intercept Coping	.15	.02	< .001
Slope Coping	.10	.04	< .05
Intercept körperliche Beschwerden	.31	.04	< .001
Slope körperliche Beschwerden	.29	.08	< .001
Intercept psychisches Wohlbefinden	.06	.02	< .01
Slope psychisches Wohlbefinden	.03	.02	= .105

Die Ergebnisse zeigen, dass sowohl die Mittelwerte der Intercepts und Slopes sowie Neurotizismus als auch die jeweiligen Varianzen signifikant sind. Es werden keinerlei cross-lag-

ged-Effekte signifikant. Die Interkorrelationen der Skalen jeweils zu einem Messzeitpunkt können dem Anhang entnommen werden (Tabelle A24, S. XLIII), es treten keine überraschenden Beziehungen auf. Überraschend ist der Befund, dass der latente Intercept-Faktor von Coping positiv mit dem latenten Slope-Faktor von psychischem Wohlbefinden zusammenhängt. Dies bedeutet, dass das psychische Wohlbefinden derjenigen Befragten, die auf einem hohen Niveau problemzentrierte Bewältigungsstrategien verwenden, über den Befragungszeitraum schneller absinkt im Vergleich zu solchen Befragten, die weniger auf problemzentrierte Strategien zurückgreifen. Zu diskutieren ist auch der direkte Pfad von dem latenten Intercept-Faktor von Coping auf den latenten Intercept-Faktor von psychischem Wohlbefinden. Psychisches Wohlbefinden wurde zum Beginn der Befragung erhoben, während Coping erst zum zweiten Messzeitpunkt mit aufgenommen wurde. Somit wäre auch ein umgekehrter Pfad, also von dem Intercept-Faktor von psychischem Wohlbefinden auf den Intercept-Faktor von Coping sinnvoll. Wird die Wirkrichtung umgekehrt, zeigt das Endmodell allerdings eine deutlich schlechtere Datenanpassung ($\chi^2 = 284.84$, $df = 173$, $p < .001$, $CFI = .95$, $RMSEA = .06$, $p\text{-close} = .13$, $\chi^2/df = 1.65$), der Pfad wird auch nicht signifikant ($\beta = -.09$, $p = .28$).

Multigruppenanalyse

Auch in den Datensätzen L1 und L3 zeigt das multivariate ALT-Modell eine gute Datenanpassung (L1: $\chi^2 = 243.71$, $df = 173$, $p < .001$, $CFI = .95$, $RMSEA = .07$, $p\text{-close} = .09$, $\chi^2/df = 1.41$; L3: $\chi^2 = 278.64$, $df = 173$, $p < .001$, $CFI = .96$, $RMSEA = .05$, $p\text{-close} = .79$, $\chi^2/df = 1.61$).

Eine Multigruppenanalyse zeigt, dass es zwischen den drei Datensätzen keinerlei Unterschiede in den Modellparametern gibt (s. Tabelle 49). Es wurden mehrere hierarchisch abgestufte Modelle getestet:

- *Modell 1.* ohne Restriktion (Alle Parameter können in beiden Gruppen variieren, Basismodell)
- *Modell 2.* Faktorladungen sind in beiden Gruppen gleich
- *Modell 3:* wie 2 und Strukturgewichte sind gleich
- *Modell 4.* wie 3 und Messintercepts sind gleich
- *Modell 5.* wie 4 und Varianzen der latenten Faktoren sind gleich
- *Modell 6.* wie 5 und Fehlervarianzen sind gleich

Tabelle 49. Ergebnisse der Multigruppenanalyse zum Vergleich des multivariaten ALT-Modells in allen drei Datensätzen

Modell 1	Modell 2	Modell 3	Modell 4	Modell 5	Modell 6
Fit-Indizes	χ^2 -Differenztest				
$\chi^2 = 791.55$, $df = 521$, $p < .001$, CFI = .96, RMSEA = .03, $p\text{-close} = 1.00$, $\chi^2/df = 1.52$	$\Delta\chi^2 = 1.96$, $df = 8$, $p = .98$	$\Delta\chi^2 = 6.92$, $df = 26$, $p = 1.00$	$\Delta\chi^2 = 9.21$, $df = 20$, $p = .98$	$\Delta\chi^2 = 11.68$, $df = 26$, $p = .99$	$\Delta\chi^2 = 22.65$, $df = 76$, $p = 1.00$

Anmerkungen: *Modell 1.* ohne Restriktion, kein Parameter wurde zwischen beiden Stichproben gleichgesetzt; *Modell 2.* Faktorladungen sind in beiden Gruppen gleich; *Modell 3.* wie 2 und Strukturgewichte sind gleich; *Modell 4:* wie 3 und Interkorrelationen sind gleich; *Modell 5.* wie 4 und Residualvarianz der Strukturvariablen ist gleich und Interkorrelationen sind gleich; *Modell 6.* wie 5 und Fehlervarianzen sind gleich

Dieses Ergebnis zeigt neben der Stabilität des ALT-Modells, dass ein möglicher selektiver Dropout keinen Einfluss auf die Ergebnisse hatte.

Die Stabilität der Parameter für das modifizierte Stressmodell wurde wie für das Messmodell für beide Stichproben mit der Bootstrap-Methode überprüft (s. Tabelle 50).

Table 50. Ergebnisse der Berechnung des multivariaten ALT-Modells mit der Bootstrap-Methode für alle Datensätze (unstandardisierte Parameter)

	Datensatz		
	L1	L2	L3
Mittlere Zunahme der Standardfehler der Faktorladungen und Strukturgewichte	39.77	19.33%	18.63%
Mittlere Zunahme der Standardfehler der Intercepts	30.63%	17.75%	15.61%
Mittlere Zunahme der Standardfehler der Kovarianzen	32.14%	19.92%	18.94%
Mittlerer Bias in den Schätzern der Faktorladungen und Strukturgewichten	.02	.01	.01
Mittlerer Bias in den Schätzern der Intercepts	.04	.03	.01
Mittlerer Bias in den Schätzern der Kovarianzen	<.01	<.01	<.01

Die Ergebnisse der Überprüfung der Messmodelle mit der Bootstrap-Methode zeigen, dass für alle Messzeitpunkte und alle Datensätze der Bias in den Parameterschätzern in einem akzeptablen Bereich liegt, der mittlere Bias liegt maximal bei $M = .02$. Darüber hinaus zeigen die Bias-korrigierten Konfidenzintervalle, dass alle signifikanten geschätzten Parameter auch hier nicht die Null einschließen. Die Verzerrungen, die möglicherweise durch die Abweichung der Daten von der multivariaten Normalverteilung verursacht sein können, fallen relativ gering aus und üben auf die Parameterschätzer nur geringen Einfluss aus. Alle Parameter, die unter Verwendung der ML-Methode statistisch signifikant sind, sind dies auch bei Verwendung der Bootstrap-Methode.

Zur weiteren Prüfung der Stabilität des ALT-Modells wurde für alle Datensätze das Modell mit latenten Indikator-Faktoren der Modellvariablen zu jedem Zeitpunkt erneut berechnet. Die Komplexität des Modells steigt hierdurch stark (311 zu schätzende Parameter im Vergleich zu 79 Parametern im nicht latenten Fall). Aufgrund der hierfür relativ kleinen

Stichprobengrößen der Datensätze im Verhältnis zu den zu schätzenden Parametern (Jackson, 2001, 2003) müssen die Ergebnisse vorsichtig interpretiert werden. Der Modell-Fit ist in allen drei Datensätzen in einem unbefriedigenden Bereich (L1: $\chi^2 = 2969.70$, $df = 1833$, $p < .001$, CFI = .77, RMSEA = .08, $p\text{-close} < .001$, $\chi^2/df = 1.62$; L2: $\chi^2 = 3058.55$, $df = 1833$, $p < .001$, CFI = .84, RMSEA = .06, $p\text{-close} < .001$, $\chi^2/df = 1.67$; L3: $\chi^2 = 3069.93$, $df = 1833$, $p < .001$, CFI = .87, RMSEA = .05, $p\text{-close} = .92$, $\chi^2/df = 1.68$), wobei mit zunehmender Stichprobengröße die Fit-Indizes besser werden. Von größerem Interesse sind mögliche, in allen drei Datensätzen vorliegende Abweichungen in der Bedeutsamkeit der Beziehungen zwischen den Modellvariablen. Es zeigt sich, dass lediglich der Pfad von dem latenten Intercept-Faktor der Ereignisseinschätzung auf den Slope-Faktor von Coping nicht signifikant wird. Alle übrigen Parameter bewähren sich zumindest in Datensatz L3, der mit $N = 300$ über die größte statistische Power verfügt.

7.2.6.3 Geschlechtseffekte

Abschließend soll geprüft werden, ob Geschlecht einen bedeutsamen Einfluss auf das ALT-Modell hat.⁵⁷ Hierzu wurde zum einen Geschlecht als weitere Variable in das ALT-Modell aufgenommen. Zum anderen wurde eine Multigruppenanalyse durchgeführt, um mögliche Unterschiede zwischen den weiblichen und männlichen Befragten in den Modellkomponenten zu analysieren. Da die Gruppe der befragten Männer sehr klein ist, wurde nur auf Datensatz L3 zurückgegriffen. Hier liegt der Anteil der männlichen Befragten bei $N = 81$, der der weiblichen Befragten bei $N = 213$.⁵⁸ Dennoch sollten die Ergebnisse der folgenden Analysen vorsichtig interpretiert werden, da die Größe der männlichen Stichprobe immer noch sehr klein ist und sich dies auf die Parameterschätzung auswirken kann. Die Aufnahme von Geschlecht als weitere Variable in das ALT-Modell zeigt signifikante Einflüsse auf einen Teil der Modellvariablen (s. Tabelle 51).

⁵⁷ Zwar konnte in Studie 1 nachgewiesen werden, dass Geschlecht keinen moderierenden Einfluss auf die Zusammenhänge zwischen den Modellvariablen hat. Dennoch wird hier ein möglicher Geschlechtseffekt analysiert, da die Veränderungen in bestimmten Modellvariablen über die Zeit zwischen den Geschlechtern unterschiedlich ausfallen können. Auf weitere Moderatoranalysen wurde aufgrund der geringen Stichprobengrößen verzichtet.

⁵⁸ Von sechs Befragten liegen inkonsistente Geschlechtsangaben über die Messzeitpunkte hinweg vor, aus diesem Grund wurden sie von den weiteren Analysen ausgeschlossen.

Tabelle 51. Einfluss von Geschlecht auf die latenten Modellvariablen des ALT-Modells

	Geschlecht ¹
	Standardisierte Parameterschätzer
Neurotizismus	-.16**
Intercept soziale Unterstützung	-.22**
Intercept Ereignisseinschätzung	.01
Intercept Selbstwirksamkeitseinschätzung	-.09
Intercept Coping	.02
Slope Coping	-.03
Intercept körperliche Beschwerden	-.16*
Slope körperliche Beschwerden	-.23*
Intercept psychisches Wohlbefinden	-.12
Slope psychisches Wohlbefinden	.26 ⁺

Anmerkungen: 1 Geschlecht ist kodiert mit 1 = weiblich und 2 = männlich; + $p < .10$; * $p < .05$

Wie zu erwarten ist, haben die weiblichen Befragten höhere Neurotizismuswerte und höhere Mittelwerte in dem latenten Intercept- und Slope-Faktor von körperlichen Beschwerden. D.h. die weiblichen Befragten geben mehr körperliche Beschwerden an und das Ausmaß an körperlichen Beschwerden steigt stärker im Vergleich zu den männlichen Befragten. Dagegen sinkt das psychische Wohlbefinden der männlichen Befragten über den Befragungszeitraum stärker als das der weiblichen Befragten, wenn auch der Geschlechtseffekt hier nur marginal signifikant wird.

Mittels einer Multigruppenanalyse soll nun geprüft werden, ob es zwischen den weiblichen und männlichen Befragten weitere Unterschiede im ALT-Modell gibt. Zunächst zeigt sich, dass das ALT-Modell auch in der Gruppe der weiblichen Befragten über befriedigende Fit-Indizes verfügt ($\chi^2 = 257.26$, $df = 173$, $p < .001$, $CFI = .95$, $RMSEA = .05$, $p\text{-close} = .60$, $\chi^2/df = 1.49$). Dagegen zeigt sich bei den männlichen Befragten, dass das ALT-Modell die Kovarianz-Matrix nicht gut repräsentiert ($\chi^2 = 356.74$, $df = 173$, $p < .001$, $CFI = .81$, $RMSEA = .12$, $p\text{-close} < .001$, $\chi^2/df = 2.06$). Die Multigruppenanalyse beinhaltet wiederum die

Berechnung von mehreren hierarchisch abgestuften Modellen, wobei ausgehend von einem Modell ohne Restriktion die Modelle immer restriktiver werden:

- *Modell 1.* ohne Restriktion (Alle Parameter können in beiden Gruppen variieren, Basismodell)
- *Modell 2.* Faktorladungen sind in beiden Gruppen gleich
- *Modell 3.* wie 2 und Messintercepts sind gleich
- *Modell 4.* wie 3 und Interkorrelationen sind gleich

Die Analyse zeigt zunächst, dass das Modell 1 über ausreichende Fit-Indizes verfügt ($\chi^2 = 621.16$, $df = 347$, $p < .001$, $CFI = .90$, $RMSEA = .05$, $p\text{-close} = .30$, $\chi^2/df = 1.79$). Somit stimmt die grundsätzliche Struktur des ALT-Modells zwischen beiden Stichproben überein. Allerdings zeigt der genestete Modellvergleich Abweichungen in einigen der Modellparameter. Zunächst zeigt Modell 2 eine signifikant schlechtere Anpassung an die Daten ($\Delta\chi^2 = 10.94$, $df = 4$, $p < .05$), was darauf hindeutet, dass es Unterschiede zwischen beiden Gruppen gibt. Lässt man die Ladungen des Beschwerdemaßes zu Messzeitpunkt 2 und 3 in beiden Gruppen frei schätzen, so zeigt das so modifizierte Modell 2 keine schlechtere Datenanpassung als Modell 1 ($\Delta\chi^2 = 4.48$, $df = 2$, $p < .11$). Auch in den Strukturgewichten gibt es bedeutsame Unterschiede zwischen beiden Gruppen, Modell 3 ist signifikant schlechter an die Daten angepasst als das modifizierte Modell 2 ($\Delta\chi^2 = 22.22$, $df = 13$, $p = .05$). Unterschiede sind hier auf die Pfade von Neurotizismus auf den latenten Intercept-Faktor, von körperlichen Beschwerden und jeweils den Pfaden von den latenten Intercept-Faktoren von Ereignisseinschätzung und Selbstwirksamkeitserwartungen auf den latenten Slope-Faktor von Coping zurückzuführen. Während bei den weiblichen Befragten die Pfade jeweils signifikant sind, haben die latenten Faktoren bei den männlichen Befragten keinen Einfluss. Nachdem die Restriktion für diese Pfade aufgehoben wurde, weist das so modifizierte Modell 3 keine signifikant schlechtere Datenanpassung auf als Modell 2 ($\Delta\chi^2 = 10.03$, $df = 10$, $p = .44$). Auch in den Intercepts der latenten Faktoren unterscheiden sich beide Befragtengruppen, Modell 4 ist signifikant schlechter an die Daten angepasst als Modell 3 ($\Delta\chi^2 = 35.48$, $df = 10$, $p < .001$). Wie erwartet und mit der oben aufgeführten Analyse übereinstimmend haben die weiblichen Befragten einen höheren Mittelwert in Neurotizismus. Dagegen weisen sie einen niedrigeren Mittelwert in dem latenten Intercept-Faktor von körperlichen Beschwerden auf, aber der Mittelwert in dem latenten Slope-Faktor von körperlichen Beschwerden ist höher, was bedeutet, dass über den Befragungszeitraum das Ausmaß an körperlichen Beschwerden bei Frauen

schneller steigt als bei den männlichen Befragten. Weitere bedeutsame Unterschiede bestehen in den Mittelwerten der latenten Intercept-Faktoren von sozialer Unterstützung und psychischem Wohlbefinden. Frauen verfügen über ein höheres Ausmaß an sozialer Unterstützung, dagegen geben Männer zu Erhebungsbeginn ein besseres psychisches Wohlbefinden an. Allerdings zeigt der Unterschied in dem Mittelwert des latenten Slope-Faktors von psychischem Wohlbefinden, dass über den Erhebungszeitraum das Wohlbefinden der männlichen Befragten schneller sinkt als das der weiblichen Befragten. Lässt man diese Parameter in beiden Gruppen frei variieren, so ist dieses modifizierte Modell 4 nicht mehr schlechter an die Daten angepasst als Modell 3 ($\Delta\chi^2 = 3.91$, $df = 4$, $p = .42$).

7.2.7 Zusammenfassendes Fazit der Ergebnisse aus Studie 2

Die Ergebnisse bestätigen die Struktur des Stressmodells nur in Teilen (Hypothese H1), es müssen Modifikationen vorgenommen werden und einzelne Zusammenhänge zwischen den Modellvariablen entsprechen nicht den Hypothesen. Es lassen sich keine Rückkopplungen zwischen den Modellkomponenten finden. Entgegen den Erwartungen (Hypothese H2) wirken die Veränderungen im psychischen Wohlbefinden und im Ausmaß an körperlichen Beschwerden nicht auf die Ereignis- und Ressourceneinschätzung zurück. Die Veränderungen im Bewältigungsverhalten sind bedeutungslos und haben keinen Einfluss auf das psychische Wohlbefinden und das Ausmaß an körperlichen Beschwerden. Es zeigen sich, wie erwartet, Veränderungen in den beiden Ergebnisvariablen über die Zeit. Körperliche Beschwerden steigen ausgehend von einem relativ geringen Ausgangsniveau zum ersten Messzeitpunkt über den Erhebungszeitraum, wobei zwischen t2 und t3 im Mittel keine Veränderungen auftreten. Auch das psychische Wohlbefinden der Befragten, ausgehend von einem relativ hohen Niveau, sinkt im Laufe der Datenerhebung, wobei sich auch in dieser Variable zwischen t2 und t3 im Mittel nur wenig ändert. In der bipolaren Operationalisierung von Coping mit den beiden Polen emotionszentriertes vs. problemzentriertes Bewältigungsverhalten zeigen sich leichte Schwankungen über die Zeit in dem Sinne, dass, ausgehend von einem relativ hohen Niveau problemzentrierten Bewältigungsverhaltens zu t1, die Werte im Mittel bis zum dritten Messzeitpunkt steigen, während sie zum vierten Messzeitpunkt wieder fallen. Allerdings fallen die Veränderungen gering aus und haben keine Auswirkungen auf die Ergebnisvariablen. In allen drei Variablen gibt es bedeutsame interindividuelle Unterschiede in den beschriebenen Veränderungen. Darüber hinaus zeigt sich, dass die Veränderungen in den körperlichen Beschwerden und im psychischen Wohlbefinden miteinander zusammenhängen, was sich in

der signifikanten Korrelation zwischen beiden Slope-Faktoren der Variablen ausdrückt. Stärkere Veränderungen in psychischem Wohlbefinden gehen auch mit stärkeren Veränderungen im Ausmaß an körperlichen Beschwerden einher. Keine Veränderungen lassen sich in den Modellvariablen Ausmaß an sozialer Unterstützung, Ereigniseinschätzung und Selbstwirksamkeitserwartungen feststellen. Die zum zweiten Messzeitpunkt vorhandenen interindividuellen Unterschiede bleiben über den Zeitraum der Untersuchung stabil. Gerade für die beiden Bewertungsformen wurden Veränderungen angenommen, da die transaktionale Stresstheorie rückkoppelnde Effekte des Bewältigungsverhaltens auf die Bewertung annimmt. Neurotizismus wurde nur zu t1 erhoben.

Weiterhin zeigen die Ergebnisse, dass Neurotizismus zwar einen bedeutsamen Einfluss auf die Modellvariablen mit Ausnahme von Coping hat, aber jeweils nur die Niveauunterschiede der Befragten zu Beginn der Datenerhebung erklären kann. Für keine der drei Modellvariablen, in denen eine Veränderung über die Zeit messbar ist, also Coping, körperliche Beschwerden und psychisches Wohlbefinden, lässt sich diese Veränderung mit Neurotizismus erklären. Dies spricht für das theoretische Stressmodell insofern, dass durchaus stabile interindividuelle Unterschiede im Belastungsausmaß angenommen werden, die relativ stabil sind (und die z.B. durch ein unterschiedliches Ausmaß in Neurotizismus erklärt werden können), aber dass die kognitiven Bewertungsprozesse und Bewältigungsstrategien von größerer Bedeutung für Veränderungen in den Ergebnisvariablen sind. Dieses Ergebnis relativiert die Befunde aus Studie 1, da Neurotizismus zumindest für den Erhebungszeitraum von Studie 2 die Veränderungen in den Ergebnisvariablen nicht direkt mit erklären kann. Weiterhin zeigt sich hypothesengemäß, dass Neurotizismus einen bedeutsamen Einfluss auf das Ausmaß an sozialer Unterstützung (H9), auf die Ereigniseinschätzung und die Selbstwirksamkeitserwartungen hat (H3) und wie in Studie 1 (aber entgegen den Hypothesen) das unterschiedliche Ausmaß an körperlichen Beschwerden und psychischem Wohlbefinden zu Beginn der Datenerhebung miterklärt.

Das Ausmaß an sozialer Unterstützung, welches stabil über den Befragungszeitraum bleibt, hängt, wie erwartet, negativ mit der Ereigniseinschätzung und positiv mit der Selbstwirksamkeitserwartung zusammen (H4). Je höher das Ausmaß an sozialer Unterstützung wahrgenommen wird, desto weniger bedrohlich wird das Referendariat eingeschätzt und desto höher sind die situationsspezifischen Selbstwirksamkeitserwartungen ausgeprägt.

Die kognitiven Bewertungsprozesse zeichnen sich durch Stabilität in Studie 2 aus. Sowohl die Ereigniseinschätzung als auch die Selbstwirksamkeitserwartungen bleiben über den Erhebungszeitraum stabil. In Übereinstimmung mit der Hypothese H5 beeinflussen aber

beide Modellvariablen das Bewältigungsverhalten: Eine positive Ereignisseinschätzung im Sinne einer Herausforderungseinschätzung geht mit der Nutzung von mehr problemzentrierten Bewältigungsstrategien einher (negativer Zusammenhang mit dem Intercept-Faktor von Coping), gleichzeitig führt eine solche (stabile) Situationseinschätzung dazu, dass die Schwankungen in der Nutzung problemzentrierter Bewältigungsstrategien über den Befragungszeitraum etwas stärker sind (negativer Zusammenhang mit dem Slope-Faktor von Coping). Gleichlaufende Effekte auf das Coping haben die Selbstwirksamkeitserwartungen. Konstant hohe Selbstwirksamkeitserwartungen führen dazu, dass mehr problemzentrierte Bewältigungsstrategien verwendet werden (positiver Zusammenhang mit dem Intercept-Faktor von Coping) und dies über die Zeit hinweg auch so bleibt (negativer Zusammenhang mit dem Slope-Faktor von Coping).

Keine Bestätigung, wie auch schon in Studie 1, findet Hypothese H7: Das Bewältigungsverhalten hat weder einen Einfluss auf den latenten Intercept-Faktor noch auf den latenten Slope-Faktor von körperlichen Beschwerden. Dagegen hat das Bewältigungsverhalten gemäß Hypothese H8 einen Einfluss auf das psychische Wohlbefinden. Eine höherer Rückgriff auf problemzentrierte Bewältigungsstrategien geht mit einem höheren psychischen Wohlbefinden einher (positiver Zusammenhang zwischen dem latenten Intercept-Faktor von Coping und dem latenten Intercept-Faktor von psychischem Wohlbefinden), aber gleichzeitig führt dies zu einem schnelleren Absinken des psychischen Wohlbefindens im Befragungszeitraum (positiver Zusammenhang zwischen Intercept-Faktor von Coping und Slope-Faktor von psychischem Wohlbefinden).

Es zeigt sich weiterhin ein moderierender Effekt von Geschlecht: Neben Unterschieden in den Mittelwerten der latenten Intercept-Faktoren von Neurotizismus, Ausmaß an sozialer Unterstützung, körperlichen Beschwerden und psychischem Wohlbefinden sowie den latenten Slope-Faktoren von körperlichen Beschwerden und psychischem Wohlbefinden liegt der Hauptunterschied in der Bedeutung des latenten Slope-Faktors von Coping. Bei den männlichen Befragten lässt sich keine bedeutsame Veränderung in den Bewältigungsstrategien über die Zeit feststellen. Allerdings weist der latente Slope-Faktor von Coping generell keine bedeutsamen Zusammenhänge mit den Ergebnisvariablen auf.

7.2.8 Zusatzanalysen

Ausgehend von dem Befund, dass, wie auch in Studie 1, das Ausmaß an körperlichen Beschwerden lediglich durch ein unterschiedliches Ausmaß in den Neurotizismuswerten zu t1

erklärt werden kann und die Veränderungen über den Befragungszeitraum in keinem Zusammenhang zu den Modellvariablen stehen (mit Ausnahme der Korrelation zwischen den beiden Slope-Faktoren von psychischem Wohlbefinden und Ausmaß an körperlichen Beschwerden), wird im Folgenden ein mögliches Alternativmodell diskutiert und analysiert. Es wird vermutet, dass körperliche Beschwerden eine eher distale Folge von Belastungen im Rahmen des Referendariats darstellen und durch das psychische Wohlbefinden bzw. die Veränderung im Wohlbefinden bedingt sind. Psychisches Wohlbefinden dagegen könnte eine eher proximale Ursache von referendariatsbezogenen Belastungen sein und somit direkter auf Belastungen bzw. auf einen unterschiedlichen Umgang mit den Belastungen bzw. Stressoren reagieren.

Dieser mögliche Zusammenhang zwischen psychischem Wohlbefinden und körperlichen Beschwerden soll in einer zusätzlichen Analyse geprüft werden. Basierend auf dem oben vorgestellten ALT-Modell werden beide Beeinflussungsrichtungen, also von psychischem Wohlbefinden auf körperliche Beschwerden und umgekehrt, gleichzeitig in einem nonrekursiven Modell berechnet. D.h. es werden gleichzeitig in einem Modell Pfade von beiden latenten Faktoren von psychischem Wohlbefinden auf beide latenten Faktoren von körperlichen Beschwerden und umgekehrt geschätzt.

Das Modell zeigt eine gute Anpassung an die Daten ($\chi^2 = 264.82$, $df = 169$, $p < .001$, $CFI = .96$, $RMSEA = .06$, $p\text{-close} = .26$, $\chi^2/df = 1.57$). Bei sukzessiver Entfernung nicht signifikanter Beziehungen bleibt im endgültigen Modell ein Pfad bestehen: Nur der Pfad von dem latenten Intercept-Faktor von psychischem Wohlbefinden auf den latenten Intercept-Faktor von körperlichen Beschwerden wird signifikant ($\beta = -.35$; $p < .05$), der Modell-Fit verändert sich nur marginal ($\chi^2 = 268.03$, $df = 174$, $p < .001$, $CFI = .96$, $RMSEA = .05$, $p\text{-close} = .31$, $\chi^2/df = 1.54$). Es erscheint dennoch sinnvoll, die Korrelation zwischen den Slope-Faktoren der beiden Modellvariablen mit aufzunehmen. Die Korrelation ist zwar nur marginal signifikant ($r = -.42$, $p = .08$), aber von mittlerer Höhe. Auch der Pfad von dem latenten Intercept-Faktor von psychischem Wohlbefinden auf den latenten Intercept-Faktor von körperlichen Beschwerden bleibt signifikant ($\beta = -.29$; $p = .05$), der Modell-Fit ist gut ($\chi^2 = 265.58$, $df = 173$, $p < .001$, $CFI = .96$, $RMSEA = .05$, $p\text{-close} = .33$, $\chi^2/df = 1.54$).

Zusammenfassend zeigt die Analyse, dass die mögliche alternative Beziehung zwischen psychischem Wohlbefinden und körperlichen Beschwerden keine Bestätigung in den Daten findet. Zwar wird der Pfad von dem Intercept-Faktor von psychischem Wohlbefinden auf körperliche Beschwerden signifikant, aber psychisches Wohlbefinden hat keinen direkten Einfluss auf die Veränderungen im Ausmaß körperlicher Beschwerden. Vielmehr zeigt sich

wie bereits im ursprünglichen ALT-Modell, dass die Veränderungen in beiden Variablen gemeinsam kovariieren.

7.3 *Diskussion Studie 2*

Die Ergebnisse von Studie 2 können die Struktur des theoretischen Stressmodells nur in Teilen bestätigen. Neurotizismus als Persönlichkeitsfaktor hat zwar einen direkten Einfluss auf die meisten Modellvariablen, kann die Veränderungen in diesen allerdings nicht aufklären. Veränderungen wiederum lassen sich nur im Bewältigungsverhalten, körperlichen Beschwerden und psychischem Wohlbefinden feststellen. Dagegen zeigen sich das Ausmaß an sozialer Unterstützung, die Ereigniseinschätzung und die Selbstwirksamkeitserwartungen als zeitstabil. Auch die Schwankungen im Bewältigungsverhalten weisen keine Bedeutsamkeit in der Aufklärung der Veränderungen in den Ergebnisvariablen auf. Darüber hinaus treten die zeitlichen Schwankungen nur bei weiblichen Befragten auf, bei den männlichen Befragten bleiben die interindividuellen Unterschiede in der Nutzung der Bewältigungsstrategien konstant. Die Veränderungen in körperlichen Beschwerden lassen sich nicht, wie erwartet, durch das Bewältigungsverhalten erklären. Dies korrespondiert mit Studie 1: Auch hier konnten die interindividuellen Unterschiede in körperlichen Beschwerden nur mit einem unterschiedlichen Neurotizismus-Niveau erklärt werden. Auch eine mögliche Alternativerklärung, die besagt, dass körperliche Beschwerden eine distale, über psychisches Wohlbefinden vermittelte Folge von Belastungen bzw. Bewältigungsversuchen ist, konnte nicht bestätigt werden. Die generelle Struktur des Stressmodells konnte für psychisches Wohlbefinden als eine Ergebnisvariable bestätigt werden. Persönlichkeitsfaktoren (Neurotizismus) und persönliche Ressourcen (soziale Unterstützung) beeinflussen die kognitive Bewertung (Ereigniseinschätzung und Selbstwirksamkeitserwartungen), die sich wiederum auf das Bewältigungsverhalten auswirken. Das Bewältigungsverhalten wiederum ist ein proximaler Prädiktor für das psychische Wohlbefinden.

Die konstanten interindividuellen Unterschiede in der kognitiven Bewertung und die bedeutungslosen Schwankungen im Bewältigungsverhalten deuten darauf hin, dass die befragten ReferendarInnen im Umgang mit stressigen Ereignissen im Referendariat auf Bewertungs- und Bewältigungsstile zurückgreifen (Carver & Scheier, 1994; Lazarus, 1991a). Bestimmte Formen der Bewertung und Bewältigung scheinen kognitiv leichter abrufbar zu sein (Lazarus & Launier, 1978). Es lassen sich entgegen der prozessualen Annahmen der Theorie keine Rückkopplungen zwischen dem Bewältigungsverhalten, den Bewertungsformen und den Ergebnisvariablen feststellen.

Die Stabilität in der Bewertung und Bewältigung des Referendariats kann auch durch das für die Längsschnittuntersuchung gewählte Zeitfenster verursacht sein. Möglicherweise treten Schwankungen in der täglichen Auseinandersetzung mit stressigen Ereignissen im Referendariat auf. Über längere Zeiträume hinweg gleichen sich diese Schwankungen aus, so dass im Mittel Personen Situationen und ihre Ressourcen in einer bestimmten Art und Weise bewerten und stabile Bewältigungstendenzen haben. Um diese Frage zu klären, sollten, neben der vorliegenden Untersuchung, zukünftig Mikroanalysen durchgeführt werden (Folkman, Lazarus, Dunkel-Schetter et al., 1986), die mittels vieler Messzeitpunkte, verteilt auf kürzere Zeiträume, solche (möglichen) Schwankungen eher abbilden können (z.B. Tennen, Affleck, Armeli & Carney, 2000). Allerdings zeigen die Ergebnisse, dass stabile Bewertungs- und Bewältigungsmuster für langfristige Folgen von Stress bedeutsam sind. Kurzfristige Schwankungen sind daher wohl relevanter für kurzfristige Stressergebnisse wie Emotionen oder physiologische Veränderungen.

Schließlich zeigt Studie 2 die Wichtigkeit der gewählten Analysemethode, das so genannte ALT-Modell (Bollen & Curran, 2004; Curran & Bollen, 2001), auf, welches sowohl intra- als auch interindividuelle Varianz gleichzeitig berücksichtigt und darüber hinaus komplexere Strukturen, wie im vorliegenden Fall, modellieren kann. Im Vergleich zu Studie 1 zeigt sich, dass Neurotizismus zwar interindividuelle Unterschiede in fast allen Modellvariablen mit erklären kann und somit ein wichtiges, stabiles Persönlichkeitsmerkmal darstellt, welches im Stressprozess Berücksichtigung finden muss. Allerdings hat Neurotizismus keinen direkten Einfluss auf die Veränderungen in den Ergebnisvariablen. Allein auf Basis der Ergebnisse aus Studie 1 würde man schnell zu einer andersläufigen Aussage kommen. Es bleibt zu hoffen, dass solche Analyseverfahren, insbesondere, wenn es um die Untersuchung von Prozessen geht, in Zukunft häufiger zum Einsatz kommen werden.

8 Gesamtdiskussion

Die Gesamtdiskussion untergliedert sich in drei Unterkapitel. Zunächst werden die zentralen Ergebnisse der beiden Studien zusammenfassend dargestellt und kritisch hinsichtlich der Fragestellung der Arbeit diskutiert. Auch auf Einschränkungen der Studien wird eingegangen (Kap. 8.1). Anschließend werden ein Gesamtfazit und ein Ausblick für weitere Untersuchungen gegeben (Kap. 8.2). Die Arbeit wird abgeschlossen mit einer Diskussion möglicher Implikationen der Ergebnisse der Arbeit für Präventionsmaßnahmen im Lehramtsreferendariat (Kap. 8.3).

8.1 *Die Bewährung der transaktionalen Stresstheorie*

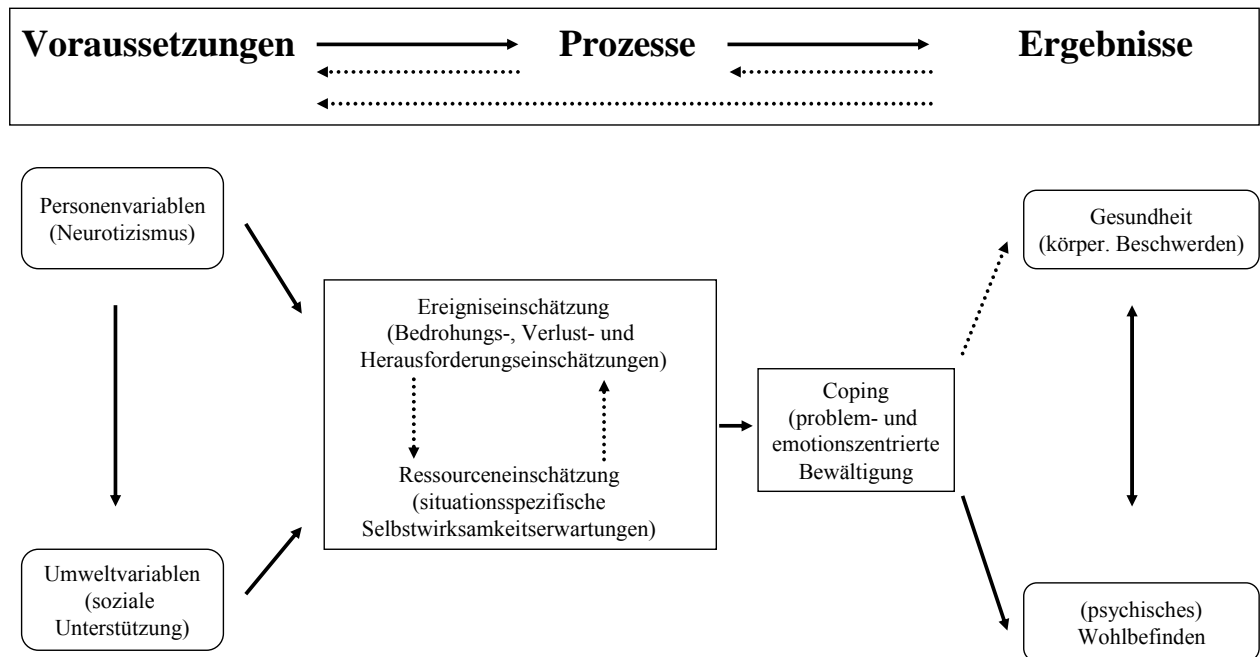
8.1.1 Zusammenfassung der Ergebnisse

Ziel der Arbeit war die Überprüfung der transaktionalen Stresstheorie von Lazarus (1991a; Lazarus & Folkman, 1984). Die Theorie definiert Stress als eine Transaktion zwischen einer Person und der Umwelt. Stress entsteht dann, wenn eine Person eine Situation als Gefahr für ihr Wohlbefinden sieht und gleichzeitig die Situation Anforderungen an die Person stellt, die ihre Ressourcen beanspruchen oder übersteigen. Kurz- und langfristige Folgen dieser Transaktion sind abhängig von den Bewältigungsstrategien, die eine Person auf Basis der Bewertung der Situation auswählt. Der Stand der empirischen Prüfung ist bislang unbefriedigend. Es finden sich eine Reihe von Untersuchungen, die direkt oder indirekt die Struktur der transaktionalen Stresstheorie überprüft haben. Gemeinsam ist allen Studien ihre unzureichende methodische Herangehensweise, die eine adäquate Überprüfung der Theorie nicht zu lassen. Hauptmängel bestehen darin, dass entweder nur interindividuelle (z.B. Chang, 1998; Jerusalem, 1993; Pakenham, 1999; Shewchuk et al., 1999; Terry, 1994) oder nur intraindividuelle Varianz (z.B. Folkman, Lazarus, Dunkel-Schetter et al., 1986; Folkman & Lazarus, 1985) bei der Auswertung berücksichtigt wurden oder wenn beide Varianzquellen gleichzeitig in die Analyse eingehen (Gunthert et al., 1999), die Struktur nicht überprüft wurde.

In zwei Studien, die im Kontext des Lehramtsreferendariats angesiedelt waren und jeweils eine Stichprobe von ReferendarInnen zu allen Komponenten der Stresstheorie befragten, wurde mittels Strukturgleichungsanalysen die transaktionale Stresstheorie auf ihre empirische Bewährung hin überprüft. Dabei sollten die oben skizzierten methodischen Mängel Berücksichtigung finden. Sowohl Struktur (Studie 1 und Studie 2) als auch der Stressprozess (Studie 2) wurden untersucht. Mit dem Stressor Lehramtsreferendariat wurde eine Situation gewählt, die für alle Befragten mehr oder weniger vergleichbar ist.

In Abbildung 10 sind die Ergebnisse beider Studien zusammenfassend dargestellt.

Abbildung 10. Zusammenfassende Darstellung der zentralen Untersuchungsergebnisse



Anmerkungen: durchgezogene Pfeile = Beziehungen konnte durch beide Studien bestätigt werden; gestrichelte Pfeile = Beziehungen konnten nicht bestätigt werden

Abbildung 10 zeigt, dass in Hinblick auf das psychische Wohlbefinden sich die Annahmen der transaktionalen Stresstheorie über die Struktur bestätigen. Sowohl in Studie 1 als auch in Studie 2 bestätigen sich die erwarteten strukturellen Beziehungen zwischen den Modellkomponenten. Personen- (Neurotizismus) und Umweltvariablen (Ausmaß an sozialer Unterstützung) beeinflussen die Ereignis- und Ressourceneinschätzung. Beide Bewertungsprozesse wirken auf das Bewältigungsverhalten, welches schließlich Varianzanteile im psychischen Wohlbefinden aufklären kann. Die zentralen Prozessvariablen haben somit den erwarteten Effekt auf das psychische Wohlbefinden und mediiieren die Effekte der Personen- und Umweltvariablen. Dagegen kann das Ausmaß an körperlichen Beschwerden nicht durch die Prozessvariablen vorhergesagt werden. Weder in Studie 1 noch in Studie 2 hat das Bewältigungsverhalten einen Einfluss auf diese Ergebnisvariable.

Auch rekursive Beziehungen zwischen den Modellkomponenten konnten nicht gefunden werden. In Studie 2 lassen sich keine rekursiven Beziehungen nachweisen, die Prozessvariablen sind über alle Messzeitpunkte stabil und spiegeln somit nur interindividuelle Unter-

schiede in den Bewertungsformen und dem Bewältigungsverhalten⁵⁹ wieder. Die prozessualen Annahmen im Sinne einer Person-Umwelt-Relation können in der vorliegenden Arbeit nicht bestätigt werden. Das Bewältigungsverhalten bewirkt nicht eine Neueinschätzung der Situation. Die Person-Umwelt-Beziehung scheint stabil über die Zeit zu sein. Die transaktionale Sichtweise von Stress hätte eine Änderung in den Bewertungsprozessen und dem Bewältigungsverhalten erwarten lassen.

8.1.2 Diskussion der Befunde

In Studie 1 konnte die Struktur der transaktionalen Stresstheorie nach einer strengen Auslegung der Ergebnisse nicht bestätigt werden. Es mussten Modifikationen vorgenommen werden, um einen akzeptablen Modell-Fit erreichen zu können. In Übereinstimmung mit den Annahmen den strukturellen Annahmen der transaktionalen Stresstheorie sind die Zusammenhänge von Neurotizismus und dem Ausmaß an sozialer Unterstützung mit der Ereignis- und Ressourceneinschätzung. Beide kognitive Bewertungsformen wirken auf das Bewältigungsverhalten, welches zumindest auf das psychische Wohlbefinden einen Einfluss hat. Allerdings zeigen die Ergebnisse aus Studie 1 auch Abweichungen, die nicht im Einklang mit den Annahmen der Theorie stehen. Neurotizismus hat, neben dem indirekten Effekt auf das Stressergebnis über die kognitive Bewertung und das Bewältigungsverhalten, auch direkte Effekte (Mak, Blewitt & Heaven, 2004; McCrae, 1990; McCrae & Costa, 1986). Somit wird der Einfluss von Neurotizismus auf das Wohlbefinden und körperliche Beschwerden nicht alleine über die Prozessvariablen der Stresstheorie mediiert. Eine mögliche Erklärung liefert Larsen (1992), der zeigen kann, dass der hohe Zusammenhang zwischen Neurotizismus und psychischen und somatischen Symptomen zum Teil auf Verzerrungen in der Enkodier- und Erinnerungsphase rückführbar ist und hierdurch Personen mit hohen Neurotizismuswerten, unabhängig von der Bewertung von und dem Umgang mit stressigen Situationen, ihre Beschwerden als schwerer erinnern als Personen mit niedrigen Neurotizismuswerten. Auch beeinflusst das Bewältigungsverhalten nicht das Ausmaß an körperlichen Beschwerden, ähnliche Ergebnisse finden Folkman, Lazarus, Gruen und DeLongis (1986), in deren Untersuchung das Bewältigungsverhalten ebenfalls keinen Einfluss auf körperliche Symptome hatte.

Die Ereignis- und Ressourceneinschätzungen interagieren nicht miteinander wie von Lazarus und Folkman (1984, 1987) angenommen. Situationsspezifische Selbstwirksamkeits-

⁵⁹ Es zeigt sich zwar in Studie 2 eine leichte zeitliche Schwankung im Bewältigungsverhalten, die auf die Gruppe der weiblichen Befragten zurückgeht. Diese zeitliche Schwankung hat aber keinen Einfluss auf die übrigen Modellvariablen und ist somit bedeutungslos.

erwartungen moderieren nicht den Einfluss der Ereigniseinschätzung auf das Bewältigungsverhalten. Vielmehr weisen beide Bewertungsformen jeweils eigenständige Effekte auf das Bewältigungsverhalten aus und sind demnach additiv, nicht multiplikativ miteinander verknüpft.

Dennoch kann als Fazit aus Studie 1 gezogen werden, dass sich an einer Stichprobe von ReferendarInnen die Struktur der transaktionalen Stresstheorie mit wenigen Modifikationen in Hinblick auf das psychische Wohlbefinden als Ergebnisvariable bestätigen lässt, wenn auch die zentralen Prozessvariablen nicht alleine die interindividuellen Unterschiede im psychischen Wohlbefinden erklären können und nicht alle Hypothesen sich bestätigen lassen.

Studie 1 weist zwei grundsätzliche Probleme auf. Zum einen lassen sich aufgrund des Querschnittsdesigns keine direkten Rückschlüsse auf kausale Beziehungen ziehen. Es wurden Strukturgleichungsanalysen als Auswertungsmethode verwendet, die es erlauben, kausale Strukturen auf ihre Passung hin auf empirisches Datenmaterial zu prüfen, aber letztlich nur als Hinweise auf die kausalen Beziehungen dienen können. Zum anderen können im Querschnitt nur interindividuelle Unterschiede untersucht werden. Intraindividuelle Veränderungen, wie es die transaktionale Stresstheorie ebenfalls postuliert, können hiermit nicht analysiert werden.

Um diesen beiden Kritikpunkten entgegen zu kommen, wurde in Studie 2 ein Längsschnittsdesign realisiert. Eine Stichprobe von ReferendarInnen wurde über einen Zeitraum von 13 Monaten zu vier Messzeitpunkten schriftlich befragt. Zur Auswertung der Längsschnittdaten wurde auf eine Auswertungsmethode zurückgegriffen, dem so genannten ALT-Modell (Bollen & Curran, 2004; Curran & Bollen, 2001), die es erlaubt, sowohl interindividuelle als auch intraindividuelle Varianz gleichzeitig zu analysieren. Solche Auswertungsverfahren wurden bislang in der Stressforschung, insbesondere zur Überprüfung der transaktionalen Stresstheorie, äußerst selten angewendet. Eine der Ausnahmen bildet eine Studie von Gunthert et al. (1999), die lineare hierarchische Verfahren (Bryk & Raudenbush, 1992) verwendeten, die es ebenfalls ermöglichen, inter- und intraindividuelle Varianz gleichzeitig zu analysieren. Im Gegensatz zu dem hier verwendeten ALT-Modell können allerdings keine strukturellen Beziehungen zwischen den Variablen analysiert werden.

Die Ergebnisse von Studie 2 bestätigen mit Abstrichen die Struktur der transaktionalen Stresstheorie und in Teilen die Ergebnisse aus Studie 1. Zunächst zeigt sich, dass einerseits das psychische Wohlbefinden im Erhebungszeitraum absinkt und andererseits das Ausmaß an körperlichen Beschwerden ansteigt. Auch im Bewältigungsverhalten gibt es Schwankungen, die aber nur bei den weiblichen Befragten festzustellen sind und darüber hinaus auch keinen

prädiktiven Wert für das Stressergebnis haben. Das Ausmaß an sozialer Unterstützung, die Ereignisseinschätzung und die situationsspezifischen Selbstwirksamkeitserwartungen bleiben über den Erhebungszeitraum stabil. Zusammengefasst lassen sich für die zentralen Prozessvariablen der transaktionalen Stresstheorie, die kognitive Bewertung (Ereignis- und Ressourceneinschätzung) und das Bewältigungsverhalten, keine Veränderungen feststellen, die von Bedeutung wären. Somit scheinen zumindest für einen länger andauernden Stressor wie das Lehramtsreferendariat Personen auf eher stabile Bewertungs- (Lazarus, 1991a) und Bewältigungsstile (Carver & Scheier, 1994; Carver et al., 1989) zurückzugreifen, mit denen sie immer wiederkehrende, ähnliche Ereignisse (im Falle des Referendariats z.B. eigenverantwortlicher Unterricht, Disziplinprobleme bei den zu unterrichtenden SchülerInnen) begegnen. Dies spricht zunächst gegen die transaktionale Stresstheorie, die von Änderungen im Laufe der Zeit ausgeht, da durch das Bewältigungsverhalten die Person-Umwelt-Relation verändert wird, was wiederum auf die Bewertung und die Bewältigung rückwirken sollte. Auch wenn sich eine solche Veränderung nicht feststellen lässt, haben die stabilen Bewertungs- und Bewältigungsstile einen Einfluss auf das Wohlbefinden. Auch die Struktur ist gemäß der Theorie: Der Bewertungsstil beeinflusst den Bewältigungsstil, der wiederum auf das Wohlbefinden wirkt. Dieser Befund spricht dafür, dass bei relativ konstanten Situationen keine oder nur geringe Veränderungen in der Bewertung und Bewältigung dieser Situationen auftreten, während möglicherweise Veränderungen über verschiedene Situationen hinweg eher zu erwarten wären (Patterson, Smith, Grant, Clopton, Josepho & Yager, 1990; Terry, 1994; s. dazu auch die Unterscheidung zwischen temporaler Konsistenz und cross-situationaler Konsistenz von Laux & Weber, 1987). In der vorliegenden Arbeit gab es keine situationale Varianz, mit dem Lehramtsreferendariat wurde ein konstanter (Makro)Stressor als Befragungskontext ausgewählt. Das Ergebnismuster, also die intraindividuelle Stabilität, kann aber auch so erklärt werden, dass durchaus kurzfristige Schwankungen vorkommen können (z.B. bei kleineren, alltäglichen Ereignissen während des Referendariats), die sich aber über einen längeren Zeitraum ausgleichen und letztlich zu dem gefundenen stabilen Muster führen. Solche kurzfristigen Muster können allerdings nur in Längsschnittuntersuchungen gefunden werden, die z.B. über einen gewissen Zeitraum die täglichen Bewertungs- und Bewältigungsprozesse untersuchen (Gunthert et al., 1999; Tennen et al., 2000).

Unterschiede und auch Veränderungen im psychischen Wohlbefinden lassen sich durch interindividuelle Unterschiede im Bewältigungsverhalten erklären. Interessant hierbei ist, dass problemzentriertes Bewältigungsverhalten mit einem höheren Wohlbefinden zu Beginn des Referendariats einhergeht, aber sich im Verlauf des Referendariats als maladaptiv herausstellt.

Die Nutzung von problemzentrierten Bewältigungsstrategien führt zu einem schnelleren Absinken des psychischen Wohlbefindens im Verlauf des Referendariats. Möglicherweise ist dies damit zu erklären, dass es sich bei einem Großteil der belastenden Situationen, mit denen man im Referendariat konfrontiert ist, um Ereignisse handelt, die man nur bedingt selbst kontrollieren und beeinflussen kann (wie z.B. das Ausmaß an eigenverantwortlichem Unterricht, die bevorstehenden Examina). Nach der transaktionalen Stresstheorie (Folkman, 1984; Folkman et al., 1979) sind problemzentrierte Bewältigungsstrategien nur dann adaptiv, wenn eine stressige Situation kontrollierbar ist (goodness-of-fit-Modell der Bewältigungseffektivität; s. Kap. 3.2). Sollte die Kontrolle nicht bei einer Person liegen, dann sind eher emotionszentrierte Strategien effektiv. Empirische Untersuchungen können die Moderatorfunktion (wahrgenommener Kontrolle) zum Teil zeigen (Forsythe & Compas, 1987; Terry & Hynes, 1998; Vitaliano et al., 1990), andere Studien finden dagegen keine Belege hierfür (Felton & Revenson, 1984; Conway & Terry, 1992; Masel et al., 1996). Die stabilen Bewertungsprozesse beeinflussen wiederum das Bewältigungsverhalten. Die konstante Bewertung des Referendariats als Bedrohung und Verlust (Carver & Scheier, 1994; Folkman & Lazarus, 1985) wie auch konstant geringe situationsspezifische Selbstwirksamkeitserwartungen (Ashford, 1988; Terry, 1994; Terry et al., 1995) führen dazu, dass eher emotionszentrierte Bewältigungsstrategien ausgewählt werden. Personale (in der vorliegenden Arbeit Neurotizismus) und situationale Ressourcen (Ausmaß an sozialer Unterstützung) beeinflussen wie angenommen die Bewertungsstile. Niedrige Neurotizismuswerte (Bolger & Zuckerman, 1995; Gunthert et al., 1999) und ein höheres Ausmaß an sozialer Unterstützung (DeLongis, Folkman & Lazarus, 1988) führen zu einer generellen Bewertung des Referendariats als Herausforderung und zu höheren Selbstwirksamkeitserwartungen. Zusammen genommen stehen diese Zusammenhänge im Einklang mit der transaktionalen Stresstheorie. Neurotizismus hat zwar auf alle übrigen Variablen, bis auf das Bewältigungsverhalten, einen Einfluss, was man, wie die direkten Einflüsse von Neurotizismus auf das psychische Wohlbefinden und das Ausmaß an körperlichen Beschwerden in Studie 1, so interpretieren kann, dass Personen mit hohen Neurotizismuswerten Informationen verzerrt (negativ) erinnern (Larsen, 1992). Aber Neurotizismus hat keinen Einfluss auf die Veränderungen im psychischen Wohlbefinden und dem Ausmaß an körperlichen Beschwerden, was einerseits für die transaktionale Stresstheorie spricht, da nur das Bewältigungsverhalten mit Veränderungen im psychischen Wohlbefinden steht, und andererseits gegen die direkten Effekte von Neurotizismus spricht, die in Studie 1 gefunden wurden. Hierauf wird später noch näher eingegangen werden.

Beide Studien weisen Stärken auf. Einerseits liegen die Stärken in der adäquaten Anwendung der gewählten Auswertungsmethoden, die bislang, zumindest in der Stressforschung, oft fehlerhaft verwendet wurden (Roesch, 1999). Es wurden jeweils die Voraussetzungen für die Methoden kritisch überprüft und für fehlende Werte adäquate Schätzverfahren verwendet. In Studie 1 wurden alternative Schätzverfahren für fehlende Werte verwendet und die Ergebnisse durch die Aufteilung der Gesamtstichprobe in zwei Substichproben kreuzvalidiert. Für die Auswertung der Studie 2 wurde auf ein aktuelles Auswertungsverfahren zurückgegriffen, welches erlaubt, sowohl intra- als auch interindividuelle Varianz gleichzeitig zu berücksichtigen als auch strukturelle Zusammenhänge zwischen Variablen zu analysieren. Eine weitere Stärke der Arbeit ist, dass alle Theorieelemente operationalisiert wurden und somit eine umfassende Theorieprüfung möglich war. Die Berücksichtigung dieser beiden Punkte, adäquate Anwendung von Auswertungsmethoden und umfassende Operationalisierung der Theorie, sprechen für die gefundenen Ergebnisse der beiden Studien, die somit die Forschung zur transaktionalen Stresstheorie bereichern. Dennoch weisen beide Studien auch Schwachstellen auf. Dies betrifft die Erhebungsmethode, den gewählten Erhebungszeitraum in Studie 2 und die Operationalisierung bestimmter Modellvariablen. Generelles Problem von Selbstberichtsdaten, auf die beide Studien zurückgreifen, ist die Gefahr von so genannter „common method variance“ (gemeinsame Varianz, die auf die gleiche Erhebungsmethode zurückgeht), die die Beziehungen zwischen den erhobenen Variablen künstlich erhöhen und zu artifiziellen Ergebnissen führen kann. Diese Kritik kann nicht vollends ausgeräumt werden, wünschenswert sind immer Daten, die auch aus anderen Quellen kommen (Beobachtungsdaten, objektive Daten zum Gesundheitszustand). Gegen diese Kritik kann angeführt werden, dass in Studie 2 zu allen realisierten Zeitpunkten die Beziehungen zwischen den Variablen gleich sind, zumal Erinnerungseinflüsse durch den langen Zeitraum zwischen den Messzeitpunkten wohl ausgeschlossen werden können. Dagegen spricht auch die bedeutsame Veränderung in nur zwei Skalen in Studie 2 (psychisches Wohlbefinden und körperliche Beschwerden). Sollte die gleiche Erhebungsmethode hauptverantwortlich für die Varianz in den Daten sein, so müsste das Muster der Veränderung in allen Variablen gleich sein. Ein weiteres Problem der Erhebungsmethode ist die häufig angeführte Kritik in der Stressforschung, dass Prädiktoren (wie z.B. das Bewältigungsverhalten) und zentrale abhängige Variablen mit ähnlichen Indikatoren erhoben werden, wodurch ebenfalls artifizielle Zusammenhänge erzeugt werden können (Coyne & Racioppo, 2000; Lazarus, DeLongis, Folkman & Gruen, 1985). In beiden Studien kann bei der Überprüfung der Messmodelle mittels konfirmatorischer Faktorenanalysen gezeigt werden, dass sich alle erhobenen Theoriekomponenten trennen lassen. Dies spricht

gegen eine mögliche Konfundierung der Variablen. Allerdings wäre es auch hier wünschenswert, Daten aus verschiedenen Quellen zu erheben. Nur hiermit würde sich dieses Problem gänzlich ausschließen lassen.

Wie bereits oben diskutiert, können mit dem gewählten Erhebungszeitraum und den Abständen zwischen den realisierten Messzeitpunkten kurzfristige Stressepisoden und ihre Bewertungs- und Bewältigungsprozesse nicht abgebildet werden. Dies ist möglicherweise der Grund dafür, dass keine Veränderungen in den meisten Modellvariablen gefunden werden konnten. Allerdings waren im Fokus der Arbeit mit psychischem Wohlbefinden und dem Ausmaß an körperlichen Beschwerden langfristige Stressergebnisse. Die Operationalisierung einiger Variablen ist auch optimierbar. Das Bewältigungsverhalten wurde mit einer Kurzsкала erfasst, die es nicht ermöglicht, verschiedene Subdimensionen zu untersuchen. Auch konnten nicht, wie erwartet, zwei Faktoren extrahiert werden, sondern lediglich eine bipolare Skala gebildet werden mit den Endpolen emotionszentrierte vs. problemzentrierte Bewältigungsstrategien. Auch für die Ereignisbewertung ließ sich nur ein Faktor extrahieren. Weitere Untersuchungen sollten diese Schwachstellen aufheben.

8.2 Abschließendes Fazit und Ausblick

Die Ergebnisse aus beiden Studien zusammen genommen können nur Teile der Vorhersagen der transaktionalen Stresstheorie von Lazarus (Lazarus, 1991a; Lazarus & Folkman, 1984) bestätigen. Die strukturellen Annahmen werden durch beide Studien bedingt unterstützt. Personale Ressourcen haben Effekte auf die Ereignis- und Ressourceneinschätzung, die wiederum das Bewältigungsverhalten beeinflussen. Zumindest in Form des psychischen Wohlbefindens kann auch ein Effekt für das Bewältigungsverhalten auf das Stressergebnis festgestellt werden. Gegen die Stresskonzeption der transaktionalen Stresstheorie sprechen dagegen zwei Befunde. In der Längsschnittstudie lassen sich keine Veränderungen über die Zeit für die kognitive Bewertung und nur bedeutungslose Schwankungen für das Bewältigungsverhalten feststellen. Dies spricht eher für stabile Bewertungs- und Bewältigungsmuster, die Personen in länger andauernden Stressepisoden, wie das Lehramtsreferendariat, nutzen (vgl. Carver & Scheier, 1994; Carver et al., 1989; Lazarus, 1991a). Damit verbunden zeigen sich keine rekursiven Effekte. Die transaktionale Stresstheorie geht davon aus, dass der Umgang mit einer stressigen Situation (und die Folgen daraus) auf die Situation und deren Bewertung rückwirkt. Die Person-Umwelt-Relation soll sich über die Zeit ändern. Es konnten in Studie 2 keine rekursiven Effekte gefunden werden. Darüber hinaus zeigen sich nur wenige cross-lagged Effekte, also Beziehungen zwischen Variablen über die Zeit. Ein dynamisch

verlaufender Stressprozess lässt sich in der vorliegenden Arbeit nicht finden. Vielmehr deuten die Ergebnisse aus Studie 2 darauf hin, dass Personen im Umgang mit länger andauernden stressigen Situationen wie dem Lehramtsreferendariat auf stabile, vermutlich gelernte Bewertungsmuster zurückgreifen und auch die Bewältigungsversuche konstant bleiben.

Durch die Prozessvariablen der Stresstheorie können weder interindividuelle Unterschiede zwischen den Befragten noch intraindividuelle Veränderungen im Ausmaß an körperlichen Beschwerden erklärt werden. Dies spricht ebenfalls gegen die transaktionale Stresstheorie. Möglicherweise beeinflussen die objektiven Belastungsfaktoren wie z.B. Unterrichten, Vorbereitung auf die Examina und den Unterricht, Umgang mit SchülerInnen, direkt das körperliche Wohlbefinden. Kognitive Bewertungsprozesse und ein unterschiedlicher Umgang mit solchen Belastungsfaktoren scheinen das Auftreten oder das Ausbleiben von körperlichen Beschwerden nicht zu beeinflussen. Eine weitere, eher statistische Erklärung steht in Zusammenhang mit der Auswahl des Lehramtsreferendariats als Stressor. Möglicherweise hätte eine Variation des Stressors, also mehrere unterschiedliche stressige Situationen, mehr Varianz in den Prozessvariablen und im Ausmaß an körperlichen Beschwerden erzeugt und einen Effekt der Prozessvariablen auf das Ausmaß an körperlichen Beschwerden wahrscheinlicher gemacht. Zukünftige Untersuchungen können hierüber Aufschluss geben.

Zuletzt stellt sich die Frage, ob die transaktionale Stresstheorie mit ihren strukturellen und prozessualen Annahmen überhaupt prüfbar ist. Schwarzer (1998a) kommt zu dem Schluss, dass nur Ausschnitte der Theorie analysiert werden können, während die generellen Annahmen zu Struktur und zum Prozess nicht testbar wären. Neue Analysemethoden, wie das in der vorliegende Arbeit eingesetzte ALT-Modell (Bollen & Curran, 2004; Curran & Bollen, 2001), ermöglichen aber eine umfassende und gleichzeitige Analyse von sowohl strukturellen Beziehungen zwischen Variablen wie auch Veränderungen über die Zeit. Aber auch mit solchen komplexen Methoden scheint eine Falsifizierbarkeit der Theorie nur schwer möglich. Finden sich die angenommenen Beziehungen nicht, lassen sich leicht Argumente für einen inadäquaten methodischen Zugang finden wie z.B. ein falscher Erhebungszeitraum, unvalide Operationalisierungen. Auch für die vorliegende Arbeit können solche Einschränkungen angeführt werden. So wurden in Studie 2 vier Messzeitpunkte realisiert, die relativ weit auseinander lagen, kurzfristige Prozesse wurden nicht analysiert. Wünschenswert ist, dass in zukünftigen Untersuchungen verstärkt stabile und variable Formen der Bewertung und der Bewältigung stressiger Situationen analysiert werden. Dabei sollte die Beziehung der Variablen über einen längeren Zeitraum und gleichzeitig über kurze Zeiträume beobachtet werden (Lazarus, 2000; Tennen et al., 2000). Der Aufwand solcher Untersuchungen ist sicherlich

groß, doch nur so kann das Zusammenspiel zwischen stabilen und dynamischen Strukturen verstanden werden.

Schon Lazarus konstatierte Anfang der 90er Jahre des letzten Jahrhunderts (1993b; S. 236-237), dass der relative Beitrag von Dispositionen und kontextualen Prozessen nicht verstanden werden kann, „...unless coping strategies are examined in the same person over time or across stressful encounters“ (Lazarus, 1993b, pp. 236-237). Wenige Arbeiten liegen dazu vor, nur wenige Studien verwenden angemessene Auswertungsmethoden. Von unmittelbarer Wichtigkeit sind Analyseverfahren wie das in dieser Arbeit eingesetzte ALT-Modell (Bollen & Curran, 2004; Curran & Bollen, 2001), die es ermöglichen, komplexe Beziehungen zwischen Variablen wie auch die Veränderung über die Zeit zu modellieren. Die Anwendung solcher Verfahren wird die weitere Theoriebildung in der Stressforschung vorantreiben. Die vorliegende Arbeit soll ein erster Schritt in diese Richtung sein.

8.3 Implikationen für eine theoriegeleitete Stressprävention

Welche Implikationen haben die gefundenen Ergebnisse für Präventionsmaßnahmen im Lehramtsreferendariat? Zunächst zeigen die Befunde, dass Prävention indiziert ist. Bereits in den ersten Monaten des Lehramtsreferendariats sinkt das psychische Wohlbefinden und das Ausmaß an körperlichen Beschwerden steigt. Diese Veränderungen treffen nicht auf alle ReferendarInnen gleichermaßen zu, es bestehen substanzielle interindividuelle Unterschiede. Dies deutet darauf hin, dass mit geeigneten Maßnahmen versucht werden kann, den ReferendarInnen Kompetenzen zu vermitteln, besser mit den Belastungen im Referendariat umzugehen.

Die transaktionale Stresstheorie liefert Hinweise darauf, wodurch diese interindividuellen Unterschiede zu Stande kommen. Genau hier können Interventionsmaßnahmen ansetzen. Dies zeigt gleichzeitig, wie wichtig auf Theorien basierende Forschung ist. Befunde aus Untersuchungen können direkt für die Planung von Präventionsmaßnahmen genutzt werden, da sie Wissen über Ursache-Wirkungs-Zusammenhänge liefern und Hinweise geben, an welchen Variablen sinnvoll angesetzt werden kann. Hierdurch ist der Erfolg von Maßnahmen wahrscheinlicher, aber natürlich nicht gewährleistet. Wie kann nun eine Präventionsmaßnahme basierend auf den vorliegenden Untersuchungsergebnissen aussehen?

Eine Möglichkeit besteht darin, auf personale und situationale Ressourcen Einfluss zu nehmen. Gerade Neurotizismus hat als relativ stabiles Persönlichkeitsmerkmal Einfluss auf eine Reihe von kritischen Stressvariablen. Eine direkte Beeinflussung scheint allerdings schwierig, wobei es generell Hinweise darauf gibt, dass Neurotizismus sich im Laufe des

Lebens verringert (Srivastava, John, Gosling & Potter, 2003). Dennoch können durch geeignete Screening-Verfahren Personen identifiziert werden, die aufgrund hoher Neurotizismuswerte eine erhöhte Wahrscheinlichkeit aufweisen, mit den Belastungen im Referendariat nur ungenügend umzugehen. Neurotizismus könnte demnach als Diagnostikum dienen, um entscheiden zu können, wer eine Intervention besonders benötigt. Allerdings wird es schwierig sein, einen Cut-off Wert zu finden. Aus diesem Grund sollte, wenn ausreichend finanzielle Mittel zur Verfügung stehen, geeignete Präventionsmaßnahmen allen ReferendarInnen angeboten werden. Auch das Ausmaß an sozialer Unterstützung wird sich nur schwer direkt beeinflussen lassen.

Von größerer Bedeutung für Präventionsmaßnahmen scheinen die zentralen Prozessvariablen der transaktionalen Stresstheorie zu sein: Die kognitive Bewertung und das Bewältigungsverhalten (Christ, Wagner & van Dick, 2002, 2004). Durch gezielte Maßnahmen kann hier direkt Einfluss genommen werden. Eine Möglichkeit bieten Stressbewältigungstrainings, die sich allgemein in verschiedenen Anwendungsbereichen als effektiv erwiesen haben (s. hierzu Meta-Analysen von Bamberg & Busch, 1996; Kaluza, 1997). Auch für LehrerInnen existieren bereits evaluierte Maßnahmen (z.B. Bertochi, Elwin, Curley & Borg, 1989; Cecil & Forman, 1990; Jenkins & Calhoun, 1991; Kramis-Aebischer, 1995; Long, 1988; Rudow, 1996), die sich als effektiv erwiesen haben, viele Trainings haben (neben zum Teil weiteren Theorien) die transaktionale Stresstheorie zur Grundlage (z.B. das leider noch nicht evaluierte Stressmanagementtraining für LehrerInnen von Kretschmann, 2000, 2004).

An dieser Stelle soll auf das Stressbewältigungstraining von Kaluza (1996, 2004) ausführlicher eingegangen werden. Es handelt sich dabei um ein primärpräventives Programm, welches auf der transaktionalen Stresstheorie basiert und gezielt auf die kognitive Bewertung und das Bewältigungsverhalten Einfluss nimmt. Es ist als Gruppentraining konzipiert und bietet somit einen ökonomischen Einsatz, eine Gruppe sollte zwischen sechs bis maximal 12 Personen umfassen. Insgesamt erstreckt sich das Programm auf 12 zweistündige Sitzungen und umfasst drei zentrale Bausteine: Ein Problemlösetraining, ein Genusstraining und ein Entspannungstraining. Mit Hilfe dieser Bausteine soll das Bewältigungsrepertoire der TeilnehmerInnen erweitert werden und stressbezogene Bewertungsprozesse verändert werden. Wichtig ist, dass der Autor zu recht betont (1996, S. 58), dass nicht bestimmte Bewältigungsstrategien vermittelt werden sollen, sondern dass die TeilnehmerInnen flexibel in stressigen Situationen unterschiedliche Bewältigungsstrategien, ob problemzentriert oder emotionszentriert, anwenden können. Dies ist auch in Übereinstimmung mit den Ergebnissen der vorliegenden Untersuchung: Problemzentrierte Bewältigungsstrategien gehen zwar mit einem bes-

serem psychischen Wohlbefinden einher, führen aber im Laufe des Referendariats dazu, dass das Wohlbefinden schneller sinkt! Das Training zeigt in einer sorgsam durchgeführten Evaluationsstudie die angenommenen Effekte (Kaluza, 1998, 1999), sowohl kurz- als auch langfristig: Die TrainingsteilnehmerInnen zeigen im Vergleich zu einer Kontrollgruppe ein größeres Bewältigungsrepertoire und ein besseres psychisches Befinden.

Das Stressbewältigungstraining sollte nicht eins zu eins bei ReferendarInnen eingesetzt werden. Ein größerer Wirkungsgrad ist anzunehmen, wenn das Training an die besondere Situation im Lehramtsreferendariat angepasst ist. Gerade der Baustein „Problemlösetraining“ bietet sich hierfür an. Hier werden, an von den TeilnehmerInnen berichteten, alltäglichen Situationen neue Wege der Bewältigung in Rollenspielen geübt. Eine Fokussierung auf wichtige referendariatspezifische Situationen scheint sinnvoll.

Wünschenswert vor der breiten Anwendung solcher Stressbewältigungstrainings ist weitere Forschung im Kontext des Lehramtsreferendariats, die folgend kurz umrissen werden soll. Die Anwendung der transaktionalen Stresstheorie hat sich, was die vorliegende Arbeit zeigen kann, im Kontext des Lehramtsreferendariats bewährt und mit dem allgemeinen Stressbewältigungsprogramm von Kaluza (1996, 2004) liegt eine primärpräventive Maßnahme vor, die dezidiert auf der transaktionalen Stresstheorie basiert und bereits evaluiert ist. Somit scheint die Durchführung von Trainings, die sich an dem Training orientieren, sinnvoll und förderungswert. Dennoch scheinen noch weitere Schritte vorab empfehlenswert. Untersuchungen sind wünschenswert, die in Form von Mikroanalysen versuchen, in Abhängigkeit von spezifischen Situationen, mit denen ReferendarInnen häufig konfrontiert sind, adaptive Bewältigungsstrategien zu identifizieren. Berücksichtigung sollte auch der normative Kontext finden, der einen Einfluss darauf hat, welche Strategien vom Umfeld gewünscht (und gefordert) sind und welche eher sanktioniert werden (Weber, 1997). Die Ergebnisse solcher Studien sollten unmittelbar in ein referendariatspezifisches Stressbewältigungstraining einfließen. Ein weiterer Schritt, der wichtig erscheint, ist, nach der situationsspezifischen Entwicklung und vor dem breiten Einsatz eines solchen Stressbewältigungstrainings, eine umfassende und sorgsame Evaluation dieses Trainings (für einen Überblick über Evaluationen s. Hager, Patry & Brezing, 2000; Wottawa & Thierau, 1998; für spezielle methodische Gesichtspunkte Kaluza & Schulze, 2000). Erst nach dem sich eine Maßnahme als effektiv in der Anwendung im Lehramtsreferendariat erwiesen hat, sollte es angeboten werden. Wenn es sich allerdings als effektiv erweisen sollte, dann scheint der breite Einsatz mehr als wünschenswert. Nicht nur die ReferendarInnen selbst würden profitieren. Auch das Schulsystem wie auch alle weiteren Personen, die unmittelbar in Kontakt mit den ReferendarInnen stehen, würden indirekt

einen Vorteil aus solchen Maßnahmen ziehen. Das System Schule braucht leistungsfähige (zukünftige) LehrerInnen. SchülerInnen und AusbilderInnen werden Unterschiede zwischen einer belasteten und erschöpften oder einer motivierten und optimistischen Referendarin erkennen (s. Heyse, 2004; Sieland, 2004). Somit würden Stresspräventionsmaßnahmen eine sinnvolle Ausbildungsergänzung darstellen, eine Integration in das Ausbildungscurriculum scheint einfach, das oben beschriebene Stressbewältigungstraining von Kaluza (1996, 2004) nimmt 12 zweistündige Sitzungen in Anspruch.

Schließlich soll neben dieser individualpsychologischen Perspektive auf Prävention nicht das System selbst vergessen werden. Rudow (1994, 2000) stellt arbeits- und organisationsbezogene Maßnahmen vor, die Belastungen im Lehrerberuf allgemein vorbeugen können. Aber auch im Referendariat selbst existieren systembedingte Faktoren, die verändert werden können, um das dem Referendariat inhärente Stresspotenzial abzubauen. So sollte der Leistungsdruck reduziert werden, der durch ein hohes Lehrdeputat und vor allem durch die Verkopplung von Beratung und Benotung bei Unterrichtsbesuchen erzeugt wird (Oesterreich, 1987; Ulich, 1996; Wissenschaftsrat, 2001). Ein weiteres strukturelles Problem, welches viel diskutiert ist, stellt die geringe Verzahnung zwischen erster (universitäre Ausbildung) und zweiter Ausbildungsphase (Referendariat) dar (Sieland, 2004). Deutlich soll hiermit gemacht werden, dass auch das System große Verantwortung für die Belastungen der ReferendarInnen trägt. Eine effektive Prävention muss sowohl bei der Person als auch am System ansetzen.

9 Literatur

- Admiraal, W. F., Korthagen, F. A. J. & Wubbels, T. (2000). Effects of student teachers' coping behavior. *British Journal of Educational Psychology*, 70, 33-52.
- Affleck, G., Tennen, H., Urrows, S. & Higgins, P. (1994). Person and contextual features of daily stress reactivity: Individual differences in relations of undesirable daily events with mood disturbance and chronic pain intensity. *Journal of Personality and Social Psychology*, 66, 329-340.
- Aiken, L. S. & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Aitken, J. L. & Mildon, D. (1991). The dynamics of personal knowledge and teacher education. *Curriculum Inquiry*, 21, 141-162.
- Aldwin, C. M. & Revenson, T. A. (1987). Does coping help? A reexamination of the relation between coping and mental health. *Journal of Personality and Social Psychology*, 53, 337-348.
- Amirkhan, J. H. (1990). A factor analytically derived measure of coping: The Coping Strategy Indicator. *Journal of Personality and Social Psychology*, 59, 1066-1074.
- Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling* (pp. 243-277). Mahwah, NJ: Erlbaum.
- Arbuckle, J. L. (2003). *Amos 5 user's guide supplement*. Chicago, IL: SmallWaters Corporation.
- Arbuckle, J. L. & Wothke, W. (1999). *Amos 4.0 user's guide*. Chicago, IL: SmallWaters Corporation.
- Arnold, M. B. (1960). *Emotion and personality* (2nd ed.). New York: Columbia University Press.
- Ashford, S. J. (1988). Individual strategies for coping with stress during organizational transitions. *Journal of Applied Behavioral Science*, 24, 19-36.
- Aspinwall, L. G. & Taylor, S. E. (1992). Modeling cognitive adaptation: A longitudinal investigation of the impact of individual differences and coping on college adjustment and performance. *Journal of Personality and Social Psychology*, 63, 989-1003.

- Averill, J. R. & Rosenn, M. (1972). Vigilant and non-vigilant coping strategies and psychophysiological stress reactions during anticipation of electric shock. *Journal of Personality and Social Psychology*, 23, 128-141.
- Bagozzi, R. P. & Edwards, J. R. (1998). A general approach for representing constructs in organizational research. *Organizational Research Methods*, 1, 45-87.
- Bagozzi, R. P. & Heatherton, T. F. (1994). A general approach for representing multifaceted personality constructs: Application to state self-esteem. *Structural Equation Modeling*, 1, 35-67.
- Baltes, P. B. & Nesselroade, J. R. (1979). History and rationale of longitudinal research. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 1-39). New York: Academic Press.
- Bamberg, E. & Busch, C. (1996). Betriebliche Gesundheitsförderung durch Stressmanagementtraining: Eine Metaanalyse (quasi)experimenteller Studien. *Zeitschrift für Arbeits- und Organisationspsychologie*, 40, 127-137.
- 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. (2002). Item parceling issues in structural equation modeling. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 269-296). Mahwah, NJ: Erlbaum.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84, 191-215.
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, 37, 122-147.
- Baron, R. M. & Kenny, D. A. (1986). The moderator-mediator distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.
- Beckmann, D., Brähler, E. & Richter, H.-E. (1991). *Der Gießen-Test (GT)* (4. erweit. und überarb. Auflage mit Neustandardisierung). Bern: Huber.
- Ben-Porath, Y. S. & Tellegen, A. (1990). A place for traits in stress research. *Psychological Inquiry*, 1, 14-17.
- 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, CA: Multivariate Software.
- Bentler, P. M. & Yuan K.-H. (1999). Structural equation modeling with small samples: Test statistics. *Multivariate Behavioral Research*, 34, 181-197.
- Bhagat, R. S. & Allie, S. M. (1989). Organizational stress, personal life stress, and symptoms of life strains: An examination of the moderating role of sense of competence. *Journal of Vocational Behavior*, 35, 231-253.
- Billings, A. G. & Moos, R. H. (1984). Coping, stress, and social resources among adults with unipolar depression. *Journal of Personality and Social Psychology*, 46, 877-891.
- Bjorck, J. P. & Cohen, L. H. (1993). Coping with threats, losses, and challenges. *Journal of Social and Clinical Psychology*, 12, 56-72.
- Blascovich, J. & Tomaka, J. (1996). The biopsychosocial model of arousal regulation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 1-51). New York: Academic Press.
- Böhmman, M. (2000). Grenzerfahrung Referendariat - Wege, Tipps, Auswege. In P. Daschner & U. Drews (Hrsg.), *Kursbuch Referendariat* (3. Auflage, S. 52-56). Weinheim: Beltz.
- Böhmman, M. (2001). *Fit für das Referendariat. Der Dschungelführer für Schule und Seminar*. Lichtenau: AOL.
- Bolger, N. (1990). Coping as a personality process: A prospective study. *Journal of Personality and Social Psychology*, 59, 525-537.
- Bolger, N. & Eckenrode, J. (1991). Social relationships, personality, and anxiety during a major stressful event. *Journal of Personality and Social Psychology*, 61, 440-449.
- Bolger, N. & Zuckerman, A. (1995). A framework for studying personality in the stress process. *Journal of Personality and Social Psychology*, 69, 890-902.
- Bollen, K. A. (1987). Outliers and improper solutions: A confirmatory factor analysis example. *Sociological Methods and Research*, 15, 375-384.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bollen, K. A. & Curran, P. J. (2004). Autoregressive latent trajectory (ALT) models: A synthesis of two traditions. *Sociological Methods & Research*, 32, 336-383.
- Borkenau, P. & Ostendorf, F. (1991). Ein Fragebogen zur Erfassung fünf robuster Persönlichkeitsfaktoren. *Diagnostica*, 37, 29-41.

- Borkenau, P. & Ostendorf, F. (1993). *NEO-Fünf-Faktoren Inventar (NEO-FFI) - Handanweisung*. Göttingen: Hogrefe.
- Boyle, G. J., Borg, M. G., Falzon, J. M. & Baglioni, A. J. Jr. (1995). A structural model of the dimensions of teacher stress. *British Journal of Educational Psychology*, 65, 49-67.
- Brähler, E. & Scheer, J. W. (1983). *Der Gießener Beschwerdefragebogen (GBB)*. Bern: Huber.
- Brähler, E., Schumacher, J. & Brähler, C. (2000). Erste gesamtdeutsche Normierung der Kurzform des Gießener Beschwerdeboogens. *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 50, 14-21.
- Brandstädter, J. (1989). Personal self-regulation and development: Cross-sectional analyses of development-related control beliefs and emotions. *Developmental Psychology*, 25, 96-108.
- Brief, A. P., Burke, M. J., George, J. M., Robinson, B. S. & Webster, J. (1988). Should negative affectivity remain an unmeasured variable in the study of job stress. *Journal of Applied Psychology*, 73, 193-198.
- Brown, C. H. (1983). Asymptotic comparison of missing data procedures for estimating factor loadings. *Psychometrika*, 48, 269-291.
- Brown, R. L. (1994). Efficacy of the indirect approach for estimating structural equation model with missing data: A comparison of five methods. *Structural Equation Modeling*, 1, 287-316.
- Browne, M. W. (1982). Covariance structures. In D. M. Hawkins (Ed.), *Topics in applied multivariate analysis* (pp. 72-141). Cambridge, UK: Cambridge University.
- Browne, M. W. (1984). Asymptotically distribution-free methods for analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37, 62-83.
- Browne, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.
- Browne, M. W. & Du Toit, S. H. C. (1991). Models for learning data. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change* (pp. 47-68). Washington, DC: American Psychological Association.
- Bryk, A. S. & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.

- Bullinger, M. (1995). German translation and psychometric testing of the SF-36 Health Survey: Preliminary results from the IQOLA project. *Social Science & Medicine*, *41*, 1359-1366.
- Bullinger, M. & Kirchberger, I. (1998). *SF-36: Fragebogen zum Gesundheitszustand* (Handanweisung). Göttingen: Hogrefe.
- Bullinger, M., Kirchberger, I. & Ware, J. E. (1995). Der deutsche SF-36 Health Survey: Übersetzung und psychometrische Testung eines krankheitsübergreifenden Instruments zur Erfassung der gesundheitsbezogenen Lebensqualität. *Zeitschrift für Gesundheitswissenschaften*, *3*, 21-36.
- Burke, M. J., Brief, A. P. & George, J. M. (1993). The role of negative affectivity in understanding relations between self-reports of stressors and strains: A comment on the applied psychology literature. *Journal of Applied Psychology*, *78*, 402-412.
- Burke, R. J. & Greenglass, E. (1993). Work stress, role conflict, social support, and psychological burnout among teachers. *Psychological Reports*, *73*, 371-380.
- Burke, R. J., Greenglass, E. R. & Schwarzer, R. (1996). Predicting teacher burnout over time: Effects of work stress, social support, and self-doubts on burnout and its consequences. *Anxiety, Stress and Coping*, *9*, 261-275.
- Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications and programming*. Mahwah, NJ: Erlbaum.
- Byrne, B. M., Shavelson, R. J. & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*, 456-466.
- Byrne, D. (1964). Repression-Sensitization as a dimension of personality. In B. A. Maher (Ed.), *Progress in experimental personality research* (Vol. 1, pp. 169-220). New York: Academic Press.
- Capel, S. A. (1991). A longitudinal study of burnout in teachers. *British Journal of Educational Psychology*, *61*, 36-45.
- Carver, C. S., Pozo, C., Harris, S. D., Noriega, V., Scheier, M. F., Robinson, D. S., Ketcham, A. S., Moffat, F. L., Jr. & Clark, K. C. (1993). How coping mediates the effects of optimism on distress: A study of women with early stage breast cancer. *Journal of Personality and Social Psychology*, *65*, 375-391.

- Carver, C. S. & Scheier, M. F. (1990). Principles of self-regulation: Action and emotion. In P. C. Kendall (Ed.), *Advances in cognitive-behavioral research and therapy* (Vol. 2, pp. 3-52). San Diego, CA: Academic Press.
- Carver, C. S. & Scheier, M. F. (1994). Situational coping and coping dispositions in a stressful transaction. *Journal of Personality and Social Psychology*, 6, 184-195.
- Carver, C. S., Scheier, M. F. & Weintraub, J. K. (1989). Assessing coping strategies: A theoretically based approach. *Journal of Personality and Social Psychology*, 56, 267-283.
- Cecil, M. A. & Forman, S. G. (1990). Effects of stress inoculation training and coworker support groups on teacher's stress. *Journal of School Psychology*, 28, 105-118.
- Chalmers, B. E. (1981). A selective review of stress: Some cognitive approaches taken a step further. *Current Psychological Reviews*, 1, 325-344.
- Chan, D. W. (1998). Stress, coping strategies, and psychological distress among secondary school teachers in Hong Kong. *American Educational Research Journal*, 35, 145-163.
- Chan, D. W. (2002). Stress, self-efficacy, social support, and psychological distress among prospective Chinese teachers in Hong Kong. *Educational Psychology*, 22, 557-569.
- Chang, E. C. (1998). Dispositional optimism and primary and secondary appraisal of a stressor: Controlling for confounding influences and relations to coping and psychological and physical adjustment. *Journal of Personality and Social Psychology*, 74, 1109-1120.
- Chang, E. C. (2000). Causal uncertainty and depressive symptoms: Appraisals and coping as mediating variables. *Journal of Social and Clinical Psychology*, 19, 420-436.
- Chou, C.-P., Bentler, P. M. & Satorra, A. (1991). Scaled test statistics and robust standard errors for non-normal data in covariance structure analysis: A Monte Carlo study. *British Journal of Mathematical and Statistical Psychology*, 44, 347-357.
- Christ, O., Wagner, U. & van Dick, R. (2002). Krise Referendariat. *HLZ (Hessische Lehrerinnen und Lehrer Zeitung)*, 55 (9), 17-21.
- Christ, O., Wagner, U. & van Dick, R. (2004). Belastung und Beanspruchung bei Lehrerinnen und Lehrern in der Ausbildung. In A. Hillert & E. Schmitz (Hrsg.), *Psychosomatische Erkrankungen bei Lehrerinnen und Lehrern* (S. 113-119). Stuttgart: Schattauer.
- Cobb, V. L. (1999). An international comparison of teacher education. *Eric Digest*, 2.

- Cohen, F. & Lazarus, R. (1979). Coping with the stresses of illness. In G. C. Stone, F. Cohen & N. E. Adler (Eds.), *Health psychology: A handbook* (pp. 217-254). San Francisco, CA: Jossey-Bass.
- Cohen, S., Kessler, R. C. & Gordon, L. U. (1995). Strategies for measuring stress in studies of psychiatric and physical disorders. In S. Cohen, R. C. Kessler & L. U. Gordon (Eds.), *Measuring stress. A guide for health and social scientists* (pp. 3-26). New York: Oxford University Press.
- Cole, D. A. & Maxwell, S. E. (1985). Multitrait-Multimethod comparisons across populations: A confirmatory factor analytic approach. *Multivariate Behavioral Research*, 20, 389-417.
- Conway, V. J. & Terry, D. J. (1992). Appraised controllability as a moderator of the effectiveness of different coping strategies: A test of the goodness of fit hypothesis. *Australian Journal of Psychology*, 44, 1-7.
- Costa, P. T. & McCrae, R. R. (1985). *The NEO personality inventory: Manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T. & McCrae, R. R. (1989). *The NEO-PI/NEO-FFI manual supplement*. Odessa, FL: Psychological Assessment Resources.
- Cox, T. (1978). *Stress*. London: MacMillan.
- Cox, T. & Mackay, C. (1981). A transactional model to occupational stress. In E. N. Corlett & J. Richardson (Eds.), *Stress, work design, and productivity* (pp. 91-113). New York: Wiley.
- Coyne, J. C. & Racioppo, M. (2000). Never the twain shall meet? Closing the gap between coping research and clinical intervention research. *American Psychologist*, 55, 655-664.
- Crandall, R. & Perrewe, P. L. (1995). *Occupational stress: A handbook*. Philadelphia, PA: Taylor & Francis.
- Cudeck, R. & Browne, M. W. (1983). Cross-validation of covariance structures. *Multivariate Behavioral Research*, 18, 147-167.
- Curran, P. J. & Bollen, K. A. (2001). The best of both worlds: Combining autoregressive and latent curve models. In A. Sayer & L. Collins (Eds.), *New methods for the analysis of change* (pp. 107-135). Washington, DC: American Psychological Association.

- Curran, P. J. & Hussong, A. M. (2002). Structural equation modeling of repeated measures data: Latent curve analysis. In D. S. Moskowitz & Hershberger, S. L. (Eds.), *Modeling intraindividual variability with repeated measures data* (pp. 59-85). Mahwah, NJ: Erlbaum.
- Curran, P. J., West, S. G. & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*, 16-29.
- Daschner, P. & Drews, U. (Hrsg.) (2000). *Kursbuch Referendariat* (3. Aufl.). Weinheim: Beltz.
- David, J. P., Green, P. J., Martin, R. & Suls, J. (1997). Differential roles of neuroticism, extroversion and event desirability for mood in daily life: An integrative model of top-down and bottom-up influences. *Journal of Personality and Social Psychology, 72*, 149-159.
- Davidov, E., Schmidt, P. & Bamberg, S. (2004). Modeling longitudinal data of an intervention study on travel mode choice: Combining latent growth curves and autoregressive models. In K. Montfort, H. Oud & A. Satorra (Eds.), *Recent developments in structural equation models: Theory and application* (pp. 263-294). Dordrecht, NL: Kluwer.
- DeLongis, A., Folkman, S. & Lazarus, R. S. (1988). The impact of daily stress on health and mood: Psychological and social resources as mediators. *Journal of Personality and Social Psychology, 54*, 486-495.
- Dewe, P. J. (1985). Coping with work stress: An investigation of teachers' actions. *Research in Education, 33*, 27-40.
- Dewe, P. J. (1991). Primary appraisal, secondary appraisal and coping: Their role in stressful work encounters. *Journal of Occupational Psychology, 64*, 331-351.
- Dohrenwend, B. S. & Dohrenwend, B. P. (Eds.) (1974). *Stressful life events: Their nature and effects*. Oxford: Wiley.
- Drumheller, P. M., Eicke, F. J. & Scherer, R. F. (1991). Cognitive appraisal and coping of students varying in stress level during three stages of a college examination. *Journal of Social Behavior and Personality, 6*, 237-254.
- Duncan, S. C. & Duncan, T. E. (1995). Modeling the processes of development via latent variable growth curve methodology. *Structural Equation Modeling, 2*, 187-213.

- Dunkley, D. M., Zuroff, D. C. & Blankstein, K. R. (2003). Self-critical perfectionism and daily affect: Dispositional and situational influences on stress and coping. *Journal of Personality and Social Psychology*, 84, 234-252.
- Durdel, A. (2000). Was tut sich wo - Ein Blick von Bundesland zu Bundesland. In P. Daschner & U. Drews (Hrsg.), *Kursbuch Referendariat* (3. Aufl., S. 196-215). Weinheim: Beltz.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7, 1-26.
- Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans*. Philadelphia, PA: SIAM.
- Efron, B. & Tibshirani, R. J. (1993). *An introduction to bootstrap*. New York: Chapman and Hall.
- Elkerton, C. A. (1984). An investigation of stress experienced by beginning teachers during practice teaching. *South African Journal of Education*, 4, 97-102.
- Enders, C. K. (2001). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling*, 8, 128-141.
- Enders, C. K. & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*, 8, 430-457.
- Enders, C. K. & Peugh, J. L. (2004). Using a EM covariance matrix to estimate structural equation models with missing data: Choosing an adjusted sample size to improve the accuracy of inferences. *Structural Equation Modeling*, 11, 1-19.
- Endler, N. S. & Parker, J. D. (1990). Multidimensional assessment of coping: A critical evaluation. *Journal of Personality and Social Psychology*, 58, 844-854.
- Expertengruppe Lehrerbildung (2002). *Empfehlungen zur Aktualisierung der Lehrerbildung in Hessen*. Wiesbaden: Hessisches Kultusministerium.
- Felton, B. J. & Revenson, T. A. (1984). Coping with chronic illness: A study of illness controllability and the influence of coping strategies on psychological adjustment. *Journal of Consulting and Clinical Psychology*, 52, 343-353.
- Felton, B. J., Revenson, T. A. & Hinrichsen, G. A. (1984). Stress and coping in the explanation of psychological adjustment among chronically ill adults. *Social Science & Medicine*, 18, 889-898.

- Filipp, S.-H. (Hrsg.). (1995). *Kritische Lebensereignisse* (3. Aufl.). Weinheim: PVU.
- Fletcher, B. C. & Payne, R. L. (1982). Levels of reported stressors and strains amongst schoolteachers: Some UK data. *Educational Review*, 34, 267-278.
- Florian, V., Mikulincer, M. & Taubman, O. (1995). Does hardiness contribute to mental health during a stressful real-life situation? The roles of appraisal and coping. *Journal of Personality and Social Psychology*, 68, 687-695.
- Folkman, S. (1984). Personal control and stress and coping processes: A theoretical analysis. *Journal of Personality and Social Psychology*, 46, 839-852.
- Folkman, S. & Lazarus, R. S. (1980). An analysis of coping in a middle-aged community sample. *Journal of Health and Social Behavior*, 21, 219-239.
- Folkman, S. & Lazarus, R. S. (1981). Reply to Shinn and Krantz. *Journal of Health and Social Behavior*, 22, 457-459.
- Folkman, S. & Lazarus, R. S. (1985). If it changes it must be a process: Study of emotion and coping during three stages of a college examination. *Journal of Personality and Social Psychology*, 48, 150-170.
- Folkman, S., Lazarus, R. S., Dunkel-Schetter, C., DeLongis, A. & Gruen, R. J. (1986). Dynamics of a stressful encounter: Cognitive appraisal, coping, and encounter outcomes. *Journal of Personality and Social Psychology*, 50, 992-1003.
- Folkman, S., Lazarus, R. S., Gruen, R. J. & DeLongis, A. (1986). Appraisal, coping, health status, and psychological symptoms. *Journal of Personality and Social Psychology*, 50, 571-579.
- Folkman, S. & Moskowitz, J. T. (2004). Coping: Pitfalls and promises. *Annual Review of Psychology*, 55, 745-774.
- Folkman, S., Schaefer, C. & Lazarus, R.S. (1979). Cognitive processes as mediators of stress and coping. In V. Hamilton & D. Warburton (Eds.), *Human stress and cognition* (pp. 265-298). Chichester, UK: Wiley.
- Forsythe, C.J. & Compas, B.E. (1987). Interaction of cognitive appraisals of stressful events and coping: Testing the goodness of fit hypothesis. *Cognitive Therapy and Research*, 11, 473-485.
- Fritzsche, K. P. (1998). *Die Stressgesellschaft*. München: Kösel.
- Fugate, M., Kinicki, A. J. & Scheck, C. L. (2002). Coping with an organizational merger over four stages. *Personnel Psychology*, 55, 905-928.

- Gallagher, D. J. (1990). Extraversion, neuroticism and appraisal of stressful academic events. *Personality and Individual Differences, 10*, 1053-1057.
- Garst, H., Frese, M. & Molenaar, P. C. M. (2000). The temporal factor in stressor-strain relationships: A growth curve model on a longitudinal study in East Germany. *Journal of Applied Psychology, 85*, 417-438.
- Glyshwa, K., Cohen, L. H. & Towbes, L. C. (1989). Coping strategies and psychological distress: Prospective analyses of early and middle adolescents. *American Journal of Community Psychology, 17*, 607-623.
- Gorrell, J. J., Bregman, N. J., McAllister, H. A. & Lipscomb. (1985). An analysis of perceived stress in elementary and secondary student teachers and full-time teachers. *Journal of Experimental Education, 54*, 11-14.
- Gowan, M. A., Riordan, C. M. & Gatewood, R. D. (1999). Test of a model of coping with involuntary job loss following a company closing. *Journal of Applied Psychology, 84*, 75-86.
- Graham, J. W., Hofer, S. M. & MacKinnon. (1996). Maximize the usefulness of data obtained with planned missing value patterns: An application of maximum likelihood procedures. *Multivariate Behavioral Research, 31*, 197-218.
- Graham, J. W. & Schafer, J. L. (1999). On the performance of multiple imputation for multivariate data with small sample size. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (S. 1-29). Thousand Oakes, CA: Sage.
- Greenglass, E. R., Fiksenbaum, L. & Burke, R. J. (1994). The relationship between social support and burnout over time in teachers. *Journal of Social Behavior and Personality, 9*, 219-230.
- Greve, W. (1997). Sparsame Bewältigung - Perspektiven für eine ökonomische Taxonomie von Bewältigungsformen. In C. Tesch-Römer, C. Salewski & G. Schwarz (Hrsg.), *Psychologie der Bewältigung* (S. 18-41). Weinheim: PVU.
- Gunthert, K. C., Cohen, L. H. & Armeli, S. (1999). The role of neuroticism in daily stress and coping. *Journal of Personality and Social Psychology, 77*, 1087-1100.
- Guttman, L. A. (1954). A new approach to factor analysis. The radax. In P. F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences* (pp. 258-348). New York: Columbia University Press.

- Hager, W., Patry, J.-L. & Brezing, H. (2000). *Evaluation psychologischer Interventionsmaßnahmen: Standards und Kriterien*. Bern: Huber.
- Handke, U. (1997). *Der Mutmacher. Ratgeber für den pädagogischen Berufseinstieg*. Berlin: Cornelson Scriptor.
- Haney, C. J. & Long, B. C. (1995). Coping effectiveness: A path analysis of self efficacy, control, coping and performance in sport competitions. *Journal of Applied Social Psychology, 25*, 1726-1746.
- Hertzog, C. & Nesselroade, J. R. (2003). Assessing psychological change in adulthood: An overview of methodological issues. *Psychology and Aging, 18*, 639-657.
- Hessisches Statistisches Landesamt. (2001). *Der Lehrernachwuchs für das Lehramt an allgemein bildenden und an beruflichen Schulen in Hessen 2001*. Wiesbaden: Hessisches Statistisches Landesamt.
- Heyse, H. (2004). Lehrgesundheit - eine Herausforderung für Schulen und Schuladministration. In A. Hillert & E. Schmitz (Hrsg.), *Psychosomatische Erkrankungen bei Lehrerinnen und Lehrern* (S. 223-239). Stuttgart: Schattauer.
- Hillert, A. & Schmitz, E. (2004). *Psychosomatische Erkrankungen bei Lehrerinnen und Lehrern*. Stuttgart: Schattauer.
- Hobfoll, S. E. (1988). *The ecology of stress*. New York: Hemisphere Publishing Corporation.
- Hobfoll, S. E. (1989). Conversation of resources: A new attempt at conceptualizing stress. *American Psychologist, 44*, 513-524.
- Hobfoll, S. E. (1998). *Stress, culture, and community: The psychology and philosophy of stress*. New York: Plenum.
- Hobfoll, S. E. (2001). The influence of culture, community, and the nested-self in the stress process: Advancing Conversation of Resources Theory. *Applied Psychology: An International Review, 50*, 337-370.
- Hobfoll, S. E., Schwarzer, R. & Chon, K. K. (1998). Disentangling the stress labyrinth: Interpreting the meaning of the term stress as it is studied in the health context. *Anxiety, Stress and Coping, 11*, 181-212.
- Holmes, T. H. & Rahe, R. H. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research, 11*, 213-218.
- Holohan, C. J. & Moos, R. H. (1985). Life stress and health: Personality, coping, and family support in stress resistance. *Journal of Personality and Social Psychology, 49*, 739-747.

- Holohan, C. J. & Moos, R. H. (1987). Personal and contextual determinants of coping strategies. *Journal of Personality and Social Psychology*, 52, 946-955.
- Hoogland, J. J. & Boomsma, A. (1998). Robustness studies in covariance structure modeling: An overview and a meta-analysis. *Sociological Methods & Research*, 26, 329-367.
- Horn, J. L. & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, 18, 117-144.
- Houtman, I. L. & Bakker, F. C. (1991b). Individual differences in reactivity to and coping with the stress of lecturing. *Journal of Psychosomatic Research*, 35, 11-24.
- Houtman, I. L. & Bakker, F. C. (1991a). Stress and coping in lecturing, and the stability of responses across practice. *Journal of Psychosomatic Research*, 35, 323-333.
- Hu, L.-T. & Bentler, P. M. (1995). Evaluation model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues and applications* (pp. 76-99). Thousand Oakes: Sage.
- Hu, L.-T. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Hu, L.-T., Bentler, P. M. & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, 112, 351-362.
- Hudek- Knežević, J. & Kardum, I. (2000). The effects of dispositional and situational coping, perceived social support, and cognitive appraisal on immediate outcome. *European Journal of Psychological Assessment*, 16, 190-201.
- Jaccard, J. & Wan, C. K. (1995). Measurement error in the analysis of interaction effects between continuous predictors using multiple regression: Multiple indicator and structural equation approaches. *Psychological Bulletin*, 117, 348-357.
- Jackson, D. L. (2001). Sample size and number of parameter estimates in maximum likelihood confirmatory factor analysis: A Monte Carlo investigation. *Structural Equation Modeling*, 8, 205-223.
- Jackson, D. L. (2003). Revisiting sample size and number of parameter estimates: Some support for the N:q hypothesis. *Structural Equation Modeling*, 10, 128-141.
- Janke, W., Erdmann, G. & Kallus, W. (1985). *Streßverarbeitungsfragebogen (SVF)*. Göttingen: Hogrefe.

- Jehle, P. (1997). Vorzeitige Pensionierung von Lehrerinnen und Lehrern - Befunde und Desiderate der Forschung. In Becker, S., Carle, U., Döbrich, P., Hoier, H. D. & Schönwälder, H. G. (Hrsg.), *Jahrbuch für Lehrerforschung* (Bd. 1, S. 247-360). Weinheim: Beltz.
- Jenkins, S. & Calhoun, J. F. (1991). Teacher stress: Issues and intervention. *Psychology in the Schools*, 28, 60-70.
- Jerusalem, M. (1990). *Persönliche Ressourcen, Vulnerabilität und Streßerleben*. Göttingen: Hogrefe.
- Jerusalem, M. (1991). Allgemeine Selbstwirksamkeit und differentielle Streßprozesse. *Psychologische Beiträge*, 33, 388-406.
- Jerusalem, M. (1993). Personal resources, environmental constraints, and adaptational processes: The predictive power of a theoretical stress model. *Personality and Individual Differences*, 14, 15-24.
- Jerusalem, M. (1999). HBV-L: Herausforderungs-, Bedrohungs- und Verlusteinschätzung von Lehrern. In R. Schwarzer & M. Jerusalem (Hrsg.), *Skalen zur Erfassung von Lehrer- und Schülermerkmalen. Dokumentation der psychometrischen Verfahren im Rahmen der wissenschaftlichen Begleitung des Modellversuchs selbstwirksame Schulen* (S. 80-83). Berlin: Freie Universität, Institut für Arbeits-, Organisations- und Gesundheitspsychologie.
- Jöreskog, K. G. (1967). Some contributions to maximum likelihood factor analysis. *Psychometrika*, 32, 443-482.
- Jöreskog, K. G. (1979). Statistical estimation of structural models in longitudinal development investigations. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 303-352). New York: Academic Press.
- Jöreskog, K. G. & Yang, F. (1996). Nonlinear structural equation models: The Kenny-Judd model with interaction effects. In Marcoulides, G.A. & Schumacker, R.E. (Eds.), *Advanced structural equation modelling* (pp. 57-88). Mahwah, NJ: Erlbaum.
- Kahn, J. H. & Hessling, R. M. (2001). Measuring the tendency to conceal versus disclose psychological distress. *Journal of Social and Clinical Psychology*, 20, 41-65.
- Kahn, J. H., Hessling, R. M. & Russel, D. W. (2003). Social support, health, and well being among the elderly: What is the role of negative affectivity? *Personality and Individual Differences*, 35, 5-17.
- Kaluza, G. (1996). *Gelassen und sicher im Streß: Psychologisches Programm zur Gesundheitsförderung* (2. Aufl.). Berlin: Springer.

- Kaluza, G. (1997). Evaluation von Streßbewältigungstrainings in der primären Prävention - eine Meta-Analyse (quasi-)experimenteller Feldstudien. *Zeitschrift für Gesundheitspsychologie*, 5, 149-169.
- Kaluza, G. (1998). Effekte eines kognitiv-behavioralen Streßbewältigungstrainings auf Belastungen, Bewältigung und (Wohl-)befinden - eine randomisierte, kontrollierte prospektive Interventionsstudie in der primären Prävention. *Zeitschrift für Klinische Psychologie*, 27, 234-243.
- Kaluza, G. (1999). Sind die Effekte eines primärpräventiven Stressbewältigungstrainings von Dauer? Eine randomisierte, kontrollierte Follow-up Studie. *Zeitschrift für Gesundheitspsychologie*, 7, 88-95.
- Kaluza, G. (2004). *Stressbewältigung: Trainingsmanual zur psychologischen Gesundheitsförderung*. Berlin: Springer.
- Kaluza, G. & Schulze, H.-H. (2000). Evaluation von Gesundheitsförderungsprogrammen - Methodische Stolpersteine und pragmatische Empfehlungen. *Zeitschrift zur Gesundheitspsychologie*, 8, 18-24.
- Kaluza, G. & Vögele, C. (1999). Streß und Streßbewältigung. In N. Bierbaumer, H. Flor & K. Hahlweg (Hrsg.), *Enzyklopädie der Psychologie, Band 3: Verhaltensmedizin: Grundlagen und Interventionen*. Göttingen: Hogrefe.
- Kaunitz, N., Spokane, A. R., Lissitz, R. W. & Strein, W. O. (1986). Stress in student teachers: A multidimensional scaling analysis of elicited stressful situations. *Teaching & Teacher Education*, 2, 169-180.
- Kenny, D. & Judd, C. M. (1984). Estimating the nonlinear and interactive effects of latent variables. *Psychological Bulletin*, 96, 201-210.
- Kishton, J. M. & Widaman, K. F. (1994). Unidimensional versus domain representative parceling of questionnaire items: An Empirical Example. *Educational and Psychological Measurement*, 54, 757-765.
- Klauer, T., Philipp, S. H. & Ferring, D. (1989). Der "Fragebogen zur Erfassung von Formen der Krankheitsbewältigung" (FEKB): Skalenkonstruktion und erste Befunde zu Reliabilität, Validität und Stabilität. *Diagnostica*, 35, 316-335.
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford Press.

- Kline, T. J. B. & Dunn, B. (2000). Analysis of interaction terms in structural equation models: a non-technical demonstration using the deviation score approach. *Canadian Journal of Behavioural Science*, 32, 127-132.
- Kobasa, S. C., Maddi, S. R. & Kahn, S. (1982). Hardiness and health: A prospective study. *Journal of Personality and Social Psychology*, 42, 168-177.
- Körner, A., Geyer, M. & Brähler, E. (2002). Das NEO-Fünf-Faktoren Inventar (NEO-FFI): Validierung anhand einer deutschen Bevölkerungsstichprobe. *Diagnostica*, 48, 19-27.
- Krahé, B. & Herrmann, J. (2003). Verfälschungstendenzen im NEO-FFI: Eine experimentelle Überprüfung. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24, 105-117.
- Kramis-Aebischer, K. (1995). *Stress, Belastungen und Belastungsverarbeitung im Lehrerberuf*. Bern: Paul Haupt.
- Kretschmann, R. (1994). Das Gefühl, nie fertig zu werden. *Päd extra*, 22, 6-16.
- Kretschmann, R. (Hrsg.) (2000). *Stressmanagement für Lehrerinnen und Lehrer. Ein Trainingsbuch mit Kopiervorlagen*. Weinheim: Beltz.
- Kretschmann, R. (2004). Präventive Selbsthilfe von Lehrern: Stressmanagement, Zeitmanagement, berufsbezogene Supervision. In A. Hillert & E. Schmitz (Hrsg.), *Psychosomatische Erkrankungen bei Lehrerinnen und Lehrern* (S. 207-222). Stuttgart: Schattauer.
- Krohne, H. W. (1989). The concept of coping modes: Relating cognitive person variables to actual coping behavior. *Advances in Behavior Research and Therapy*, 11, 235-248.
- Krohne, H. W. (1994). Vigilance and cognitive avoidance as concepts in coping research. In H. W. Krohne (Ed.), *Attention and avoidance* (pp. 19-50). Göttingen: Hogrefe.
- Krohne, H. W., Schumacher, A. & Egloff, B. (1992). *Das Angstbewältigungsinventar (ABI)* (Mainzer Berichte zur Persönlichkeitsforschung). Mainz: Johannes-Gutenberg-Universität, Psychologisches Institut.
- Kyriacou, C. (1998). Teacher stress: past and present. In J. Dunham. & V. Varma (Eds.), *Stress in Teachers: Past, Present and Future* (pp. 1-13). London: Whurr.
- Kyriacou, C. (2001). Teacher stress: directions for future research. *Educational Review*, 53, 27-35.
- Kyriacou, C. & Pratt, J. (1985). Teacher stress and psychoneurotic symptoms. *British Journal of Educational Psychology*, 55, 61-64.
- Kyriacou, C. & Sutcliffe, J. (1978). A model of teacher stress. *Educational Studies*, 4, 1-6.

- Lange, H. (2002). PISA: Und was nun? *Zeitschrift für Erziehungswissenschaft*, 3, 455-471.
- Larsen, R. J. (1992). Neuroticism and selective encoding and recall of symptoms: Evidence from a combined concurrent-retrospective study. *Journal of Personality and Social Psychology*, 62, 480-488.
- Laux, L. (1983). Psychologische Stresskonzeptionen. In H. Thomaе (Hrsg.), *Theorien und Formen der Motivation* (S. 453-535). Göttingen: Hogrefe.
- Laux, L. & Weber, H. (1987). Person-centered coping research. *European Journal of Personality*, 1, 193-214.
- Laux, L. & Weber, H. (1990). Bewältigung von Emotionen. In K. R. Scherer (Hrsg.), *Enzyklopädie der Psychologie: Serie Motivation und Emotion. Psychologie der Emotion* (Band 3, S. 560-629). Göttingen: Hogrefe.
- Lazarus, R. S. (1966). *Psychological stress and the coping process*. New York: McGraw-Hill.
- Lazarus, R. S. (1968). Emotions and adaptation: Conceptual and empirical relations. In W. J. Arnold (Ed.), *Nebraska Symposium on Motivation* (pp. 175-266). Lincoln: University of Nebraska Press.
- Lazarus, R. S. (1990). Theory-based stress measurement. *Psychological Inquiry*, 1, 3-13.
- Lazarus, R. S. (1991a). *Emotion and adaptation*. New York, NY, USA: Oxford University Press.
- Lazarus, R. S. (1991b). Cognition and motivation in emotion. *American Psychologist*, 46, 352-367.
- Lazarus, R. S. (1991c). Progress on a cognitive-motivational-relational theory of emotion. *American Psychologist*, 46, 819-834.
- Lazarus, R. S. (1993a). From psychological stress to the emotions: A history of changing outlooks. *Annual Review of Psychology*, 44, 1-21.
- Lazarus, R. S. (1993b). Coping theory and research: Past, present and future. *Psychosomatic Medicine*, 55, 234-247.
- Lazarus, R. S. (1999). *Stress and emotion: A new synthesis*. New York: Springer.
- Lazarus, R. S. (2000). Toward better research on stress and coping. *American Psychologist*, 55, 665-673.
- Lazarus, R. S. (2001). Conservation of resources theory (COR): Little more than words masquerading as a new theory. *Applied Psychology: An International Review*, 50, 381-391.

- Lazarus, R. S., Averill, J. R. & Opton, E. M. Jr. (1970). Toward a cognitive theory of emotions. In M. Arnold (Ed.), *Feelings and emotions* (pp. 207-232). New York: Academic.
- Lazarus, R. S. & DeLongis, A. (1983). Psychological stress and coping in aging. *American Psychologist*, *38*, 245-254.
- Lazarus, R. S., DeLongis, A., Folkman, S. & Gruen, R. (1985). Stress and adaptational outcome: The problem of confounded measures. *American Psychologist*, *40*, 770-779.
- Lazarus, R. S. & Folkman, S. (1984). *Stress, appraisal and coping*. New York: Springer.
- Lazarus, R. S. & Folkman, S. (1987). Transactional theory and research on emotions and coping. *European Journal of Personality*, *1*, 141-170.
- Lazarus, R. S. & Launier, R. (1978). Stress-related transactions between person and environment. In L. A. Pervin & M. Lewis (Eds.), *Perspectives in interactional psychology* (pp. 287-327). New York: Plenum.
- Lee, S. & Hershberger, S. (1990). A simple rule for generating equivalent models in covariance structure modeling. *Multivariate Behavioral Research*, *25*, 313-334.
- Lengua, L. J. & Long, A. C. (2002). The role of emotionality and self-regulation in the appraisal-coping process: test of direct and moderating effects. *Applied Developmental Psychology*, *23*, 471-493.
- Lenke, S. (1997). *Entwicklung und Validierung von Skalen zur sozialen Unterstützung und autoritären Tendenzen bei LehrerInnen*. Marburg: Unveröff. Diplomarbeit am Fachbereich Psychologie.
- Li, F., Harmer, P., Duncan, T. E., Duncan, S. C., Acock, A. & Boles, S. (1998). Approaches to testing interaction effects using structural equation modeling methodology. *Multivariate Behavioral Research*, *33*, 1-39.
- Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*, 1198-1202.
- Little, R. J. A. & Schenker, N. (1995). Missing data. In G. Arminger, C. C. Clog & M. E. Sobel (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (pp. 39-75). New York: Plenum.
- Little, R. J. & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Little, T. D., Cunningham, W. A., Shahar, G. & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighting the merits. *Structural Equation Modeling*, *9*, 151-173.

- Little, T. D., Lindenberger, U. & Maier, H. (2000). Selectivity and generalizability in longitudinal research: On the effects of continuers and dropouts. In T. D. Little, K. U. Schnabel & J. Baumert, J. (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches and specific examples* (pp. 187-200). Mahwah, NJ: Erlbaum.
- Long, B. C. (1988). Stress management for school personnel: Stress inoculation training and exercise. *Psychology in the Schools, 25*, 314-324.
- MacCallum, R. C. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin, 100*, 107-120.
- MacCallum, R. C. (1995). Model specification: Procedures, strategies, and related issues. In R. H. Hoyle (Ed.), *Structural equation modeling* (pp. 16-36). Thousand Oakes: Sage.
- MacCallum, R. C. & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology, 51*, 201-226.
- MacCallum, R. C., Browne, M. W. & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130-149.
- MacCallum, R. C. & Hong, S. (1997). Power analysis in covariance structure modeling using CFI and AGFI. *Multivariate Behavioral Research, 32*, 195-210.
- MacCallum, R. C., Roznowski, M., Mar, C. M. & Reith J. V. (1994). Alternative strategies for cross-validation of covariance structure models. *Multivariate Behavioral Research, 29*, 1-32.
- 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., Widaman, K. F., Zhang, S. & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*, 84-99.
- MacCallum, R. D., 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.
- MacDonald, C. J. (1992). The multiplicity of factors creating stress during the teaching practicum: The student teachers' perspective. *Education, 30*, 1-11.
- MacDonald, C. J. (1993). Coping with stress during teaching practicum: the student teachers' perspective. *The Alberta Journal of Educational Research, 4*, 407-418.

- MacDonald, C. J., MacKinnon, M., Joyce, M. & Gurney, P. (1992). Institutional constraints and role expectations: Perspectives on field experiences in the on-site teacher education program. *McGill Journal of Education*, 25, 257-273.
- Maes, S., Leventhal, H. & de Ridder, D. T. D. (1996). Coping with chronic diseases. In M. Zeidner & N. S. Endler (Eds.), *Handbook of coping: theory, research, applications* (pp. 221-251). New York: Wiley.
- Mak, A. S., Blewitt, K. & Heaven, P. C. L. (2004). Gender and personality influences in adolescent threat and challenge appraisals and depressive symptoms. *Personality and Individual Differences*, 36, 1483-1496.
- Marcoulides, G. A. & Schumacker, R. E. (2001). *New developments and techniques in structural equation modeling*. Mahwah, NJ: Erlbaum.
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57, 519-530.
- Marsh, H. W., Hau, K.-T. & Wen, Z. (2004). In search for golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11, 320-341.
- Marsh, H. W. & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher-order factor models and their invariance across groups. *Psychological Bulletin*, 97, 562-582.
- Masel, C.N., Terry, D.J. & Gribble, M. (1996). The effects of coping on adjustment: Re-examining the goodness of fit model of coping effectiveness. *Anxiety, Stress, and Coping*, 9, 279-300.
- Mason, J.W. (1971). A re-evaluation of the concept of "non-specific" in stress-theory. *Journal of Psychiatric Research*, 8, 323-333.
- Mason, J.W. (1975a). Emotion as reflected in patterns of endocrine integration. In L. Levi (Ed.), *Emotions - their parameters and measurement* (pp. 147-171). New York: Raven.
- Mason, J.W. (1975b). A historical view of the "stress" field, part II. *Journal of Human Stress*, 1, 22-36.
- Mattlin, J. A., Wethington, E. & Kessler, R. C. (1990). Situational determinants of coping and coping effectiveness. *Journal of Health and Social Behavior*, 31, 103-122.

- McArdle, J. J. (1986). Latent variable growth within behavior genetic models. *Behavior Genetics, 16*, 163-200.
- McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In J. R. Nesselroade & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (2nd ed., pp. 561-614). New York: Plenum.
- McArdle, J. J. (1998). Recent trends in modeling longitudinal data by latent growth curve models. In G. Marcoulides (Ed.), *New statistical models with business and economic applications* (pp. 359-406). Mahwah, NJ: Erlbaum.
- McArdle, J. J. & Bell, R. Q. (2000). An introduction to latent growth models for developmental data analysis. In T. D. Little, K. U. Schnabel & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches, and specific examples* (pp. 69-107). Mahwah, NJ: Erlbaum.
- McCrae, R. R. (1984). Situational determinants of coping responses: Loss, threat, and challenge. *Journal of Personality and Social Psychology, 46*, 919-928.
- McCrae, R. R. (1989). Age differences and changes in the use of coping mechanisms. *Journal of Gerontology, 44*, 161-169.
- McCrae, R. R. (1990). Controlling neuroticism in the measurement of stress. *Stress Medicine, 6*, 237-241.
- McCrae, R. R. & Costa, P. T. (1986). Personality, coping, and coping effectiveness in an adult sample. *Journal of Personality, 54*, 385-405.
- McGrath, J. E. (1970). *Social and psychological factors in stress*. New York: Holt, Rinehardt & Winston.
- McHorney, C. A., Ware, J. E., Lu, J. F. R. & Sherbourne, C. D. (1994). The MOS 36-item short-form health survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Medical Care, 32*, 40-66.
- McHorney, C. A., Ware, J. E. & Raczek, A. E. (1993). The MOS 36-item short-form health survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Medical Care, 31*, 247-263.
- Meredith, W. (1964). Notes on factorial invariance. *Psychometrika, 29*, 177-185.
- Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika, 58*, 525-543.

- Meredith, W. & Horn, J. (2001). The role of factorial invariance in modeling growth and change. In L. M. Collins and Syer, A. G. (Eds.), *New methods for the analysis of change* (pp. 203-240). Washington, DC: American Psychological Association.
- Meredith, W. & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, *55*, 107-122.
- Miller, S. (1980). When is a little knowledge a dangerous thing? Coping with stressful events by monitoring vs. blunting. In S. Levine & H. Ursin (Eds.), *Coping and health* (pp. 145-169). New York: Plenum Press.
- Millsap, R. E. & Everson, H. (1991). Confirmatory measurement model comparisons using latent means. *Multivariate Behavioral Research*, *33*, 181-220.
- Monroe, S. M. & Kelley, J. M. (1997). Measurement of stress appraisal. In S. Cohen & R. C. Kessler (Eds.), *Measuring stress: A guide for health and social scientists* (pp. 122-147). London: Oxford University Press.
- Moos, R. H. (1988). *Coping Responses Inventory. Manual*. Palo Alto, CA: Stanford University and Veterans Administration Medical Centers.
- Moos, R. H. & Billings, A. G. (1982). Conceptualizing and measuring coping resources and processes. In L. Goldberger & Breznitz, S. (Eds.), *Handbook of stress: theoretical and clinical aspects* (pp. 212-230). New York: MacMillan.
- Moos, R. H., Ralph, W. & Swindle, M. (1990). Stressful life circumstances: Concepts and measures. *Stress Medicine*, *6*, 172-178.
- Morton, L.L., Vesco, R., William, N.H. & Awender, M.A. (1997). Student teacher anxieties related to class management, pedagogy, evaluation, and staff relations. *British Journal of Educational Psychology*, *67*, 69-89.
- Moskowitz, D. S. & Hershberger, S. L. (2002). *Modeling intraindividual variability with repeated measures data: Methods and application*. Mahwah, NJ: Erlbaum.
- Moulder, B. C. & Algina, J. (2002). Comparison of methods for estimating and testing latent variable interactions. *Structural Equation Modeling*, *9*, 1-19.
- Murray-Harvey, R., Slee, P. T., Lawson, M. J., Silins, H., Banfield, G. & Russel, A. (2000). Under stress: The concerns and coping strategies of teacher education students. *European Journal of Teacher Education*, *23*, 19-35.
- Muthén, B. & Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology*, *38*, 171-189.

- Nachtigall, C., Kroehne, U., Funke, F. & Steyer, R. (2003). (Why) should we use SEM? Pros and cons of structural equation modeling. *Methods of Psychological Research Online*, 8, 1-22.
- Nevitt, J. & Hancock G. R. (2001). Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. *Structural Equation Modeling*, 8, 353-377.
- Nitsch, J. R. (1981b). Stresstheoretische Modellvorstellungen. In J. R. Nitsch (Hrsg.), *Stress: Theorien, Untersuchungen, Maßnahmen* (S. 52-141). Bern: Huber.
- Nitsch, J. R. (1981a). Zur Gegenstandsbestimmung der Stressforschung. In J. R. Nitsch (Hrsg.), *Stress: Theorien, Untersuchungen, Maßnahmen* (S. 29-51). Bern: Huber.
- Nolen-Hoeksema, S. (2001). Gender differences in depression. *Current Directions in Psychology Science*, 10, 173-176.
- Oesterreich, D. (1987). Vorschläge von Berufsanfängern für Veränderungen in der Lehrerausbildung. *Zeitschrift für Pädagogik*, 33, 771-786.
- Olsson, U. H., Foss, T., Troye, S. V. & Howell, R. D. (2000). The performance of ML, GLS, and WLS estimation in structural equation modeling under conditions of misspecification and nonnormality. *Structural Equation Modeling*, 7, 557-595.
- Pakenham, K. I. (1999). Adjustment to multiple Sclerosis: Application of a stress and coping model. *Health Psychology*, 18, 383-392.
- Pakenham, K. I. & Rinaldis, M. (2001). The role of illness, resources, appraisal, and coping strategies in adjustment to HIV/AIDS: The direct and buffering effects. *Journal of Behavioral Medicine*, 24, 259-279.
- Patterson, T. L., Smith, L. W., Grant, I., Clopton, P., Josepho, S. & Yager, J. (1990). Internal vs. external determinants of coping responses to stressful life-events in elderly. *British Journal of Medical Psychology*, 63, 149-160.
- Perrez, M. (1992). Coping-Forschung auf Abwegen? In U. Gebhard (Hrsg.), *Psychologische Erkenntnisse zwischen Philosophie und Empirie* (S. 72-88). Bern: Huber.
- Pierce, C. M. B. & Molloy, G. N. (1990). Psychological and biographical differences between secondary school teachers experiencing high and low levels of burnout. *British Journal of Educational Psychology*, 60, 37-51.
- Ping, R. A. jr. (1996a). Estimating latent variable interactions and quadratics: The state of this art. *Journal of Management*, 22, 163-183.

- Ping, R. A. jr. (1996b). Latent variable interaction and quadratic effect estimation: A two-step technique using structural equation analysis. *Psychological Bulletin*, *119*, 166-175.
- Pitts, S. C., West, S. G. & Tein, J.-Y. (1996). Longitudinal measurement models in evaluation research: Examining stability and change. *Evaluation and Program Planning*, *19*, 333-350.
- Punch, K. & Tuettemann, E. (1990). Correlates of psychological distress among secondary school teachers. *British Educational Research Journal*, *16*, 369-382.
- Roesch, S. C. (1999). Modeling stress: A methodological review. *Journal of Behavioral Medicine*, *22*, 249-269.
- Rogosa, D. R. (1988). Myths about longitudinal research. In K. W. Schaie, R. T. Campbell, W. M. Meredith & S. C. Rawlings (Eds.), *Methodological issues in aging research* (pp. 171-209). New York: Springer.
- Rogosa, D. R. (1995). Myths and methods: "Myths about longitudinal research" plus supplemental questions. In J. M. Gottman (Ed.), *The analysis of change* (pp. 4-66). Mahwah, NJ: Erlbaum.
- Rost, D. H. & Schermer, F. J. (1986). Strategien der Prüfungsangstverarbeitung. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *7*, 127-149.
- Rothermund, K. & Brandtstädter, J. (1997). Entwicklung und Bewältigung: Festhalten und Preisgeben von Zielen als Formen der Bewältigung von Entwicklungsproblemen. In C. Tesch-Römer, C. Salewski & G. Schwarz (Hrsg.), *Psychologie der Bewältigung* (S. 120-133). Weinheim: PVU.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, *61*, 581-592.
- Rudow, B. (1994). *Die Arbeit des Lehrers. Zur Psychologie der Lehrertätigkeit, Lehrerbelastung und Lehrerergesundheit*. Göttingen: Verlag Hans Huber.
- Rudow, B. (1996). *Stressmanagement bei Lehrerinnen und Lehrern in Rheinland-Pfalz. Das Projekt LEHRERGESUNDHEIT (LEGU)*. Mainz: Ministerium für Bildung, Wissenschaft und Weiterbildung.
- Rudow, B. (2000). *Arbeits- und Gesundheitsschutz im Lehrerberuf*. Ludwigsburg: Süddeutscher Pädagogischer Verlag.
- Sadowski, C. J. & Blackwell, M. W. (1985). Locus of control and perceived stress among student-teachers. *Psychological Reports*, *56*, 723-726.

- Schaarschmidt, U. & Fischer, A. (1996). *AVEM - Arbeitsbezogenes Verhaltens- und Erlebensmuster* (Manual). Frankfurt: Swets.
- Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147-177.
- Scheck, C. L. & Kinicki, A. J. (2000). Identifying the antecedents of coping with an organizational acquisition: a structural assessment. *Journal of Organizational Behavior*, 21, 627-648.
- Schedensack, P. U. (1995). Referendariat als persönliche Krise? *Pädagogik*, 12, 44-48.
- Scheier, M. F. & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology*, 4, 219-247.
- Scheier, M. F. & Carver, C. S. (1992). Effects of optimism on psychological and physical well-being: Theoretical overview and empirical update. *Cognitive Therapy and Research*, 16, 201-228.
- Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of fit measures. *Methods of Psychological Research Online*, 8, 23-74.
- Scheuch, K. & Vogel, H. (1993). Prävalenz von Befunden in ausgewählten Diagnosegruppen bei Lehrern. *Soziale Präventivmedizin*, 38, 20-25.
- Schmitz, G. S. (1999). *Zur Struktur und Dynamik der Selbstwirksamkeitserwartung von Lehrern: Ein protektiver Faktor gegen Belastung und Burnout?*. Berlin: Dissertation am Fachbereich Erziehungswissenschaften der FU Berlin.
- Schmitz, G. S. & Schwarzer, R. (2000). Selbstwirksamkeitserwartung von Lehrern: Längsschnittbefunde mit einem neuen Instrument. *Zeitschrift für Pädagogische Psychologie*, 14, 12-25.
- Schnitger, C. & van Dick, R. (2000). Die Arbeitssituation von Wissenschaftlern: Subjektive Wahrnehmung von Stress und Zufriedenheit. *Report Psychologie*, 25, 660-667.
- Scholz, U., Gutierrez-Dona, B., Sud, S. & Schwarzer, R. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment*, 18, 242-251.
- Schröder, K. E. E. & Schwarzer, R. (1997). Bewältigungsressourcen. In C. Tesch-Römer, C. Salewski & G. Schwarz (Hrsg.), *Psychologie der Bewältigung* (S. 174-195). Weinheim: PVU.

- Schwarz, G., Salewski, C. & Tesch-Römer, C. (1997). Psychologie der Bewältigung - Variationen über ein altbekanntes Thema? In C. Tesch-Römer, C. Salewski & G. Schwarz (Hrsg.), *Psychologie der Bewältigung* (S. 1-6). Weinheim: PVU.
- Schwarzer, R. (1986). *Stress, Angst und Hilflosigkeit* (2. Aufl.). Stuttgart: Kohlhammer.
- Schwarzer, R. (1998a). Stress and coping from a social-cognitiv perspective. In P. Csermely (Ed.), *Stress of life: From molecules to man* (pp. 531-537). New York: New York Academy of Sciences.
- Schwarzer, R. (1998b). Stress and coping resources: Theory and review. In Schwarzer, R. (Ed.), *Advances in health psychology research* (pp. 1-24). Berlin: Freie Universität Berlin. Institut für Arbeits-, Organisations und Gesundheitspsychologie.
- Schwarzer, R. (2000). *Stress, Angst und Handlungsregulation* (4. Aufl.). Stuttgart: Kohlhammer.
- Schwarzer, R. (2001). Stress, resources, and proactive coping. *Applied Psychology: An International Review*, 50, 400-407.
- Schwarzer, R. & Jerusalem M. (Hrsg.). (1999). *Skalen zur Erfassung von Lehrer- und Schülermerkmalen*. Berlin: Freie Universität Berlin.
- Schwarzer, R. & Schwarzer, C. (1996). A critical survey of coping instruments. In M. Zeidner & N.S. Endler (Eds.), *Handbook of Coping* (pp. 1-21). New York, NY: Wiley.
- Selye, H. (1936). A syndrome produced by diverse nocuous agents. *Nature*, 138, 32.
- Selye, H. (1976a). *The stress of life*. New York: McGraw-Hill.
- Selye, H. (1976b). *Stress in health and disease*. London: Butterworth.
- Selye, H. (1991). History and present status of the stress concept. In A. Monat & R. S. Lazarus (Eds.), *Stress and coping: An anthology* (3rd ed., pp. 21-35). New York: Columbia University Press.
- Shewchuk, R. M., Elliott, T. R., MacNair-Semands, R. R. & Harkins, S. (1999). Trait influences on stress appraisal and coping: An evaluation of alternative frameworks. *Journal of Applied Social Psychology*, 29, 685-704.
- Sieland, B. (2004). Lehrerbiographien zwischen Anforderungen und Ressourcen im System Schule. In A. Hillert & E. Schmitz (Hrsg.), *Psychosomatische Erkrankungen bei Lehrerinnen und Lehrern* (S. 143-160). Stuttgart: Schattauer.
- Simberg, S. Laine, A., Sala, E. & Rönnemaa, A.-M. (2000). Prevalence of voice disorders among future teachers. *Journal of Voice*, 14, 231-235.

- Smith, C. A. & Lazarus, R. S. (1990). Emotion and adaptation. In Pervin, L. A. (Ed.), *Handbook of personality: Theory and research* (pp. 609-637). New York, NY: Guilford Press.
- Smith, C. A. & Lazarus, R. S. (1993). Appraisal components, core relational themes, and emotions. *Cognition and Emotion*, 7, 233-269.
- Smith, E., Lemke, J., Taylor, M., Kirchner, H. L. & Hoffman, H. (1998). Frequency of voice problems among teachers and other occupations. *Journal of Voice*, 12, 480-488.
- Sörbom, D. (1974). A general method for studying differences in factor means and factor structure between groups. *British Journal of Mathematical and Statistical Psychology*, 27, 229-239.
- Somerfield, M. R. & McCrae, R. R. (2000). Stress and coping research - Methodological challenges, theoretical advances, and clinical applications. *American Psychologist*, 55, 620-625.
- Spector, P. E., Zapf, D., Chen, P. Y. & Frese, M. (2000). Why negative affectivity should not be controlled in job stress research: Don't throw out the baby with the bath water. *Journal of Organizational Behavior*, 21, 79-95.
- Srivastava, S., John, O. P., Gosling, S. D. & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change. *Journal of Personality and Social Psychology*, 84, 1041-1053.
- Starnaman, S. M. & Miller, K. I. (1992). A test of a causal model of communication and burnout in the teaching profession. *Communication Education*, 41, 40-53.
- Steenkamp, J.-B. E. M. & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25, 78-90.
- Steiger, H. J. & Lind, J. C. (1980, May). *Statistically based tests for the number of common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.
- Steinmetz, H., Schmidt, P., Tina-Booh, A. & Wiczorek, S. (in press). How invariant are measures of human values? Results of a multigroup confirmatory model analysed in a population sample. *Structural Equation Modeling*.
- Stelzl, I. (1986). Changing a causal hypothesis without changing the fit: Some rules for generating equivalent path models. *Multivariate Behavioral Research*, 21, 309-331.

- Steptoe, A. & Vögele, C. (1986). Are stress responses influenced by cognitive appraisal? an experimental comparison of coping strategies. *British Journal of Psychology*, *77*, 243-255.
- Stone, A. A. & Neale, J. M. (1984). New measures of daily coping: Development and preliminary results. *Journal of Personality and Social Psychology*, *46*, 892-906.
- Sweet, L., Savoie, J. A. & Lemyre, L. (1999). Appraisals, coping, and stress in breast cancer screening: A longitudinal investigation of causal structure. *Canadian Journal of Behavioural Science*, *31*, 240-253.
- Tabachnik, B. G. & Fidell, L. S. (2001). *Using Multivariate Statistics* (4th ed.). Boston, MA: Allyn & Bacon.
- Tamres, L. K., Janicki, D. & Helgeson, V. S. (2002). Sex differences in coping behavior: A meta-analytic review and an examination of relative coping. *Personality and Social Psychology Review*, *1*, 2-30.
- Taylor, S. E. & Aspinwall, L. G. (1996). Mediating and moderating processes in psychosocial stress. In H. B. Kaplan (Ed.), *Psychosocial stress* (pp. 71-110). San Diego, CA: Academic Press.
- Tennen, H., Affleck, G., Armeli, S. & Carney, M. A. (2000). A daily process approach to coping. *American Psychologist*, *55*, 626-636.
- Terhart, E. (2000). *Perspektiven der Lehrerbildung in Deutschland. Abschlussbericht der von der Kultusministerkonferenz eingesetzten Kommission*. Weinheim: Beltz.
- Terry, D. J. (1991). Coping resources and situational appraisals as predictors of coping behavior. *Personality and Individual Differences*, *12*, 1031-1047.
- Terry, D. J. (1994). Determinants of coping: The role of stable and situational factors. *Journal of Personality and Social Psychology*, *66*, 895-910.
- Terry, D. J. & Hynes, G. J. (1998). Adjustment to low-control situation: Reexamining the role of coping responses. *Journal of Personality and Social Psychology*, *74*, 1078-1092.
- Terry, D. J., Tonge, L. & Callan, V. J. (1995). Employee adjustment to stress: The role of coping resources, situational factors, and coping responses. *Anxiety, Stress and Coping*, *8*, 1-24.
- Thomae, H. (1987). Conceptualizations of responses to stress. *European Journal of Personality*, *1*, 171-192.

- Thompson, M. S. & Cooper, C. L. (2001). A rose by any other name: A commentary on Hobfoll's Conservation of Resources Theory. *Applied Psychology: An International Review*, 50, 408-418.
- Tisak, J. & Meredith, W. (1991). Longitudinal factor analysis. In A. von Eye (Ed.), *Statistical methods in longitudinal research: Vol. 1. Principles and structuring change* (pp. 125-150). San Diego, CA: Academic Press.
- Tobin, D. L., Holroyd, K. A., Reynolds, R. V. & Wigal, J. K. (1989). The hierarchical factor structure of the Coping Strategies Inventory. *Cognitive Therapy and Research*, 13, 343-361.
- Tomaka, J., Blascovich, J., Kelsey, R. M. & Leitten, C. L. (1993). Subjective, physiological, and behavioral effects of threat and challenge appraisal. *Journal of Personality and Social Psychology*, 65, 248-260.
- Ulich, K. (1996). Lehrer/innen-Ausbildung im Urteil der Betroffenen. *Deutsche Schule*, 88, 81-97.
- Ulman, J. B. & Bentler, P. M. (2003). Structural equation modeling. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology: Research methods in psychology* (Vol. 2, pp. 607-634). New York: Wiley & Sons.
- Urban, D. (2002). Prozessanalyse im Strukturgleichungsmodell: Zur Anwendung latenter Wachstumskurvenmodelle in der Sozialisationsforschung. *ZA-Information*, 51, 6-37.
- Valentiner, D. P., Holahan, C. J. & Moos, R. H. (1994). Social Support, appraisals of event controllability, and coping: An integrative model. *Journal of Personality and Social Psychology*, 66, 1094-1102.
- van Dick, R. (1999a). *Stress und Arbeitszufriedenheit im Lehrerberuf: Eine Analyse von Belastung und Beanspruchung im Kontext sozialpsychologischer, klinisch-psychologischer und organisationspsychologischer Konzepte*. Marburg: Tectum.
- van Dick, R. (1999b). Zwischen Engagement und Resignation. *HLZ (Hessische Lehrerinnen und Lehrer Zeitung)*, 9, 34-35.
- van Dick, R. & Wagner, U. (2001). Stress and strain in teaching: a structural equation approach. *British Journal of Educational Psychology*, 71, 243-259.

- van Dick, R., Wagner, U. & Petzel, T. (1999). Arbeitsbelastung und gesundheitliche Beschwerden von Lehrerinnen und Lehrern: Einflüsse von Kontrollüberzeugungen, Mobbing und Sozialer Unterstützung. *Psychologie in Erziehung und Unterricht*, 46, 269-280.
- van Dick, R., Wagner, U., Petzel, T., Lenke, S. & Sommer, G. (1999). Arbeitsbelastung und Soziale Unterstützung: Erste Ergebnisse einer Untersuchung von Lehrerinnen und Lehrern. *Psychologie in Erziehung und Unterricht*, 46, 55-64.
- van Dick, R., Wagner, U., Stellmacher, J. & Christ, O. (2004). The utility of a broader conceptualization of organizational identification or: Which of its aspects really matters? *Journal of Occupational and Organizational Psychology*, 77, 171-191.
- Vandenberg, R. J. & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3, 4-69.
- Vitaliano, P. P., DeWolfe, D. J., Maiuro, R. D., Russo, J. & Katon, W. (1990). Appraised changeability of a stressor as a modifier of the relationship between coping and psychological distress: A test of the hypothesis of fit. *Journal of Personality and Social Psychology*, 59, 582-592.
- Vollrath, M. & Togersen, S. (2000). Personality types and coping. *Personality and Individual Differences*, 29, 367-378.
- Ware, J. E. & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical Care*, 30, 473-483.
- Watson, D. & Pennebaker, J. W. (1989). Health complaints, stress, and distress: Exploring the central role of negative affectivity. *Psychological Review*, 96, 234-254.
- Weary, G. & Edwards, J. A. (1994). Individual differences in causal uncertainty. *Journal of Personality and Social Psychology*, 67, 308-318.
- Weber, A. (2004). Krankheitsbedingte Frühpensionierungen von Lehrkräften. In A. Hillert & E. Schmitz (Hrsg.), *Psychosomatische Erkrankungen bei Lehrerinnen und Lehrern* (S. 23-38). Stuttgart: Schattauer.
- Weber, H. (1997). Zur Nützlichkeit des Bewältigungskonzeptes. In C. Tesch-Römer, C. Salewski & G. Schwarz (Hrsg.), *Psychologie der Bewältigung* (S. 7-16). Weinheim: PVU.

- Weber, H. (2003). Breaking the rules: Personal and social responses to coping norm-violations. *Anxiety, Stress and Coping*, 16, 133-153.
- Wentura, D., Greve, W. & Klauer, T. (2002). Theorien der Bewältigung. In D. Frey & M. Irle (Hrsg.), *Theorien der Sozialpsychologie. Band III* (2., vollst. überarbeitete u. erweiterte Aufl., S. 101-125). Bern: Huber.
- West, S. G., Finch, J. F. & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56-75). Thousand Oakes, CA: Sage.
- Widaman, K. F. (2000). Testing cross-group and cross-time constraints on parameters using the general linear model. In T. D. Little, K. U. Schnabel & J. Baumert (Eds.), *Modeling longitudinal and multilevel data* (pp. 163-185). Mahwah, NJ: Erlbaum.
- Willett, J. B. (1988). Questions and answers in the measurement of change. In E. Z. Rothkopf (Ed.), *Review of research in education* (Vol. 15, pp. 345-422). Washington, DC: American Educational Research Association.
- Willett, J. B. & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, 116, 363-381.
- Wills, T. A., Sandy, J. M., Yaeger, A. M., Cleary, S. D. & Shinar, O. (2001). Coping dimensions, life stress, and adolescent substance use: A latent growth analysis. *Journal of Abnormal Psychology*, 110, 309-323.
- Wissenschaftsrat. (2001). Empfehlungen zur künftigen Struktur der Lehrerbildung vom 16. November 2001. <http://www.wissenschaftsrat.de/>, Download am 09.12.2002.
- Wothke, W. (2000). Longitudinal and multigroup modeling with missing data. In T. D. Little, K. U. Schnabel & J. Baumert (Eds.), *Modeling longitudinal and multilevel data* (pp. 219-240). Mahwah, NJ: Erlbaum.
- Wottawa, H. & Thierau, H. (1998). *Lehrbuch Evaluation* (2. Aufl.). Bern: Huber.
- Yang-Wallentin, F., Schmidt, P., Davidov, E. & Bamberg, S. (2004). Is there any interaction effect between intention and perceived behavioral control? *Methods of Psychological Research Online*, 8, 127-157.
- Yung, Y.-F. & Bentler, P. M. (1996). Bootstrapping techniques in analysis of mean and covariance structures. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling: Issues and techniques* (pp. 195-226). Mahwah, NJ: Erlbaum.

10 Anhang

Inhaltsverzeichnis

Fragebogen Studie 1.....	III
Fragebogen Studie 2.....	XII
Skalen mit Itembezeichnungen	XIX
Zusätzliche Dokumentation von Ergebnissen aus Studie 1.....	XXIV
Zusätzliche Dokumentation von Ergebnissen aus Studie 2.....	XXVI

Tabellenverzeichnis

<i>Tabelle A1.</i> Originalitems und Itembezeichnung der Neurotizismusskala	XIX
<i>Tabelle A2.</i> Originalitems und Itembezeichnung der Skala zu Messung des Ausmaßes an sozialer Unterstützung	XX
<i>Tabelle A3.</i> Originalitems und Itembezeichnung der Skala zur Messung der Herausforderungs-, Bedrohungs- und Verlusteinschätzung und der Selbstwirksamkeitserwartungen (Reihenfolge wie im Fragebogen)	XXI
<i>Tabelle A4.</i> Originalitems, Itembezeichnung und Ladung der Items auf dem Copingfaktor	XXII
<i>Tabelle A5.</i> Originalitems und Itembezeichnung der Beschwerdeskala.....	XXIII
<i>Tabelle A6.</i> Originalitems und Itembezeichnung der Skala zur Messung des psychischen Wohlbefindens.....	XXIII
<i>Tabelle A7.</i> Rücklauf für alle Studienseminare in Studie 1	XXIV
<i>Tabelle A8.</i> Item-zu-Parcel-Zuordnung in Studie 1 und Studie 2.....	XXV
<i>Tabelle A9.</i> Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t1 für den Längsschnittdatensatz L1	XXVI
<i>Tabelle A10.</i> Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t2 für den Längsschnittdatensatz L1	XXVII
<i>Tabelle A11.</i> Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t3 für den Längsschnittdatensatz L1	XXVIII

<i>Tabelle A12.</i>	Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t4 für den Längsschnittdatensatz L1	XXIX
<i>Tabelle A13.</i>	Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t1 für den Längsschnittdatensatz L2	XXX
<i>Tabelle A14.</i>	Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t2 für den Längsschnittdatensatz L2	XXXI
<i>Tabelle A15.</i>	Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t3 für den Längsschnittdatensatz L2	XXXII
<i>Tabelle A16.</i>	Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t4 für den Längsschnittdatensatz L2	XXXIII
<i>Tabelle A17.</i>	Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t1 für den Längsschnittdatensatz L3	XXXIV
<i>Tabelle A18.</i>	Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t2 für den Längsschnittdatensatz L3	XXXV
<i>Tabelle A19.</i>	Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t3 für den Längsschnittdatensatz L3	XXXVI
<i>Tabelle A20.</i>	Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t4 für den Längsschnittdatensatz L3	XXXVII
<i>Tabelle A21.</i>	Interkorrelation der Skalen über alle Messzeitpunkte für den Längsschnittdatensatz L2	XXXVIII
<i>Tabelle A22.</i>	Parameterschätzer jeweils für das einfaktorielle LGC-Modell für das Ausmaß an sozialer Unterstützung, Ereignisseinschätzung und Selbstwirksamkeitserwartungen für den Längsschnittdatensatz L2	XL
<i>Tabelle A23.</i>	Parameterschätzer jeweils für das zweifaktorielle LGC-Modell für Coping, körperliche Beschwerden und psychisches Wohlbefinden für den Längsschnittdatensatz L2	XLI
<i>Tabelle A24.</i>	Interkorrelationen (standardisierte Parameter) der Skalen jeweils innerhalb eines Messzeitpunkts für das multivariate ALT-Modell für den Längsschnittdatensatz L2	XLIII

Fragebogen Studie 1

November 2001

Fragebogen zu Belastungen und Beanspruchungen von Referendarinnen und Referendaren

Fachbereich Psychologie der Philipps-Universität Marburg
AG Sozialpsychologie
(Dipl.-Psych. Oliver Christ, Prof. Dr. Ulrich Wagner, Dr. Rolf van Dick)

Zunächst einmal vielen Dank für Ihre Teilnahme an unserer Befragung. In unserer Untersuchung beschäftigen wir uns mit Belastungen und Beanspruchungen im Referendariat.

Bislang gibt es nur sehr wenige Untersuchungen zu Belastungen und Beanspruchungen im Referendariat. In unserer Studie geht es unter anderem darum herauszufinden, welche Belastungsfaktoren für Referendarinnen und Referendare wichtig sind.

Mit unserer Untersuchung wollen wir zu Ergebnissen kommen, die möglichst zuverlässig zu verallgemeinern sind. Daher ist es besonders wichtig, dass sich alle ausgewählten Referendarinnen und Referendare an der Untersuchung beteiligen. Auf der Grundlage Ihrer Erfahrungen sollen Konzepte und Maßnahmen zur Verbesserung Ihrer Ausbildungssituation entwickelt werden.

Bei standardisierten Befragungen sind Vereinfachungen unvermeidlich. Seien Sie daher bitte nachsichtig, wenn Ihnen manche Aussagen oder Fragen verkürzt erscheinen und wählen Sie die Antwortalternative, die Ihrer Meinung am besten entspricht.

Diese Untersuchung führen wir im Rahmen eines Forschungsprojekts durch. Die Ergebnisse der Befragung werden wir zum einen in Form eines Vortrages an ihrem Studienseminar vorstellen, zum anderen werden wir Ihnen auch die Möglichkeit anbieten, Ihre individuellen Ergebnisse über das Internet mittels eines Codes abzurufen. Über die Internetadresse und das genaue Vorgehen zum Abruf der Ergebnisse werden wir Sie über Aushänge an ihrem Studienseminar rechtzeitig informieren. Selbstverständlich bleiben Ihre Angaben anonym und werden vertraulich behandelt. An keiner Stelle werden von uns Namen der Personen gespeichert.

Bitte beantworten Sie uns zunächst folgende Fragen:

Datum ___/___/_____

Ihr Geschlecht weiblich männlich

Ihr Alter _____ Jahre

Ihr derzeitiger Familienstand alleinstehend mit Partner/Partnerin

Kinder keine ja: _____ (Anzahl) Alter der Kinder: _____

Welches Lehramt streben Sie an?

Grundschule Berufsschule Sonderschule
 Haupt- und Realschule Gymnasium

Welcher Schulart gehört Ihre Ausbildungsschule an?

Grundschule Hauptschule Realschule
 Gymnasium Berufsschule Sonderschule
 Gesamtschule Sek I Gesamtschule Sek II Andere: _____

Anhang

Wie hoch schätzen Sie Ihren (durchschnittlichen) zeitlichen Aufwand für Tätigkeiten, die mit dem Referendariat zusammenhängen (incl. des Unterrichts), ein?

- während der Schulzeit? _____ Zeitstunden / Woche
 - während der Schulferien? _____ Zeitstunden / Woche

Wieviel Zeit nimmt (durchschnittlich) die Vor- und Nachbereitung der Seminarveranstaltungen an Ihrem Studienseminar in Anspruch? _____ Zeitstunden / Woche

Welche Fächer unterrichten Sie schwerpunktmäßig? _____

Wie groß ist in etwa Ihre Ausbildungsschule (Anzahl der SchülerInnen)? _____

Wie groß sind die Klassen durchschnittlich, in denen Sie hauptsächlich unterrichten? _____

Wie groß ist in etwa das Kollegium an Ihrer Ausbildungsschule? _____

Schätzen Sie bitte den durchschnittlichen Ausländeranteil in den Klassen, in denen Sie unterrichten (in %): _____

An wie vielen Klassenfahrten haben Sie bislang während Ihres Referendariats teilgenommen? _____

An wie vielen Schultagen haben Sie in den letzten drei Monaten wegen Krankheit gefehlt? an _____ Tagen

Wann haben Sie mit dem Referendariat begonnen (Monat/Jahr)? _____ / _____

In welcher Ausbildungsphase befinden Sie sich? _____

Im Folgenden sehen Sie eine Liste von Arbeitsbedingungen, die im Schulalltag bzw. während des Referendariats auftreten. Bitte kreuzen Sie bei jeder Situation an, wie belastend diese Situation aktuell für Sie ist. Ein Kreuz ganz rechts würde bedeuten, dass Sie die geschilderte Situation aktuell als sehr belastend empfinden; ein Kreuz ganz links würde bedeuten, dass Sie die geschilderte Situation aktuell als überhaupt nicht belastend empfinden. Die anderen Kreise bedeuten jeweils, dass Sie die Situation als eher belastend oder eher nicht belastend empfinden.

	überhaupt nicht belastend					sehr belastend
Zu große Klassen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mangelnde Motivation bzw. Konzentrationsfähigkeit der SchülerInnen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Disziplinprobleme	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Probleme mit ausländischen SchülerInnen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ärger mit Behörden bzw. Institutionen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Probleme mit den Eltern	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fachfremder Unterrichtseinsatz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Unterschiedliche Lernvoraussetzungen der SchülerInnen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Vor- und Nachbereitung des Unterrichts	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Geringe Lernbereitschaft von SchülerInnen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verwaltungsarbeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ständige Kritik am Lehrerberuf und fehlende Anerkennung in der Öffentlichkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	überhaupt nicht belastend				sehr belastend	
Hektisches und störungsvolles Arbeitsklima an der Schule	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Unterrichtsbesuche / Lehrproben	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Konflikte mit der Ausbildungsschule	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Eigenverantwortlicher Unterricht	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Konflikte mit AusbilderInnen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pädagogische Prüfungsarbeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Examensprüfungen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Klassenfahrten	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Seminarveranstaltungen am Studienseminar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Einstellungsdruck	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Unterrichtsentwürfe	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	überhaupt nicht				sehr	
Wie sehr belasten Sie außerberufliche Anforderungen?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Bitte füllen Sie nun die folgende Tabelle aus. Bitte tragen Sie mit Zahlen von 1-6 ein, wie sehr die Aussagen für Sie zutreffen. Die 1 bedeutet „trifft gar nicht zu“, die 6 bedeutet „trifft genau zu“. Die Zahlen 2-5 entsprechen den jeweiligen Zwischenabstufungen.

Beispiel: Wenn Sie sich sehr mit Ihrer Karriere (im Sinne beruflicher Weiterqualifikation) identifizieren, tragen Sie in das entsprechende Feld eine 6 ein, identifizieren Sie sich wenig oder gar nicht mit der Schule, tragen Sie in die entsprechenden Felder eine 2 oder 1 ein usw.

<i>1 = trifft gar nicht zu</i> <i>6 = trifft genau zu</i>	Karriere	Schule	Berufsgruppe der Lehrer
Ich identifiziere mich als karriereorientierter Mensch bzw. als Mitglied meiner Schule bzw. Berufsgruppe.			
Karriereorientiert bzw. Angehöriger meiner Schule bzw. Berufsgruppe zu sein, spiegelt meine Persönlichkeit gut wieder.			
Für die Karriere bzw. in meiner Schule/meinem Beruf zu arbeiten, macht mir viel Spaß.			
Ich denke eher ungern an meine Karriere/Schule/Berufsgruppe.			
Manchmal erzähle ich lieber nicht, dass ich karriereorientiert bzw. Mitglied meiner Schule/Berufsgruppe bin.			
Die Karriere, meine Schule/Berufsgruppe wird von anderen positiv bewertet.			
Für meine Karriere/Schule/Berufsgruppe engagiere ich mich auch über das absolut Notwendige hinaus.			

Im Folgenden geht es um Ihre Beziehungen zu wichtigen Menschen wie Familienmitgliedern, Kollegen oder Freunden. Der Fragebogen enthält eine Reihe von Aussagen. Neben jeder Aussage finden Sie sechs Kästchen. Kreuzen Sie bitte das Kästchen an, das Ihrer Meinung am besten entspricht. Es gibt keine richtigen oder falschen Antworten, es kommt auf Ihr Erleben an. Kreuzen Sie am besten ohne lange zu überlegen das Kästchen an, das Ihrer ersten Einschätzung entspricht.

	trifft nicht zu					trifft genau zu
Wenn ich mal nicht weiter weiß, kann ich mir bei meinen KollegInnen an meiner Ausbildungsschule jederzeit Rat holen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
In meiner Arbeit werde ich durch die Schüler nicht unterstützt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bei Problemen in der Schule finde ich von Seiten meines/meiner Mentors/MentorIn häufig keine Unterstützung.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe kaum Freunde, mit denen ich Freud und Leid, die das das Referendariat so mit sich bringen, teilen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Von den AusbilderInnen an meinem Studienseminar werde ich in meiner Arbeit unterstützt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin leicht beunruhigt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fühle mich anderen oft unterlegen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich stark unter Stress stehe, fühle ich mich manchmal, als ob ich zusammenbräche.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fühle mich selten einsam oder traurig.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fühle mich oft angespannt und nervös.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Manchmal fühle ich mich völlig wertlos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich empfinde selten Furcht oder Angst.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich ärgere mich oft darüber, wie andere Leute mich behandeln.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Zu häufig bin ich entmutigt und will aufgeben, wenn etwas schief geht.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin selten traurig oder deprimiert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fühle mich oft hilflos und wünsche mir eine Person, die meine Probleme löst.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Manchmal war mir etwas so peinlich, dass ich mich am liebsten versteckt hätte.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wenn im Referendariat Probleme auftreten, kann es wichtig sein, ob man sich durch verschiedene Personen unterstützt fühlt. Diese Unterstützung kann zum einen ganz praktischer Art sein, z.B. in Form wertvoller Ratschläge. Es kann sich aber auch um emotionale Unterstützung handeln, wenn man z.B. merkt, dass einem Vertrauen und Anerkennung entgegen gebracht werden oder man sich einfach bei jemandem aussprechen kann. Bitte geben Sie in der untenstehenden Tabelle möglichst genau an, als **wie stark** Sie die **praktische** bzw. **emotionale** Unterstützung durch die genannten Gruppen / Personen empfinden. Dazu stehen Ihnen Angaben von 1 bis 6 zur Verfügung. Eine 1 bedeutet dabei, dass Sie von der entsprechenden Gruppe/Person "überhaupt nicht" unterstützt werden, eine 6 bedeutet, dass Sie "voll und ganz" unterstützt werden. Die Zwischenstufen bedeuten dementsprechend eher schwächere (2 und 3) oder eher stärkere (4 und 5) Unterstützung.

1 = überhaupt keine Unterstützung 6 = völlige Unterstützung	Ausbilder- Innen	Mentoren- Innen	Referendars- kollegInnen	KollegInnen an der Ausbildungs- schule	Partner oder Partnerin und Freunde	Schüler
Wie sehr werden Sie von folgende Personen praktisch unterstützt, wenn die Arbeit im Zusammenhang mit dem Referendariat schwierig wird?						
Wie sehr fühlen Sie sich von folgenden Personen emotional unterstützt, wenn Probleme im Zusammenhang mit dem Referendariat auftauchen?						

Bitte geben Sie nun an, ob die folgenden Aussagen auf Sie persönlich zutreffen.

	trifft nicht zu					trifft genau zu
Ich bin zuversichtlich, dass ich den Anforderungen im Beruf gewachsen bin.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin mir sicher, dass ich mich in Zukunft auf individuelle Probleme der Schüler noch besser einstellen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn sich Widerstände auftun, finde ich Mittel und Wege, mich durchzusetzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin bedrückt, weil ich momentan einfach nicht klarkomme.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich traue mir zu, die Schüler für neue Projekte zu begeistern.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn eine neue Sache auf mich zukommt, weiß ich, wie ich damit umgehen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich mache mir Sorgen wegen der Probleme, die auf mich zukommen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin mir sicher, dass ich auch mit den problematischen Schülern in guten Kontakt kommen kann, wenn ich mich darum bemühe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
In letzter Zeit ist alles schief gelaufen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	trifft nicht zu					trifft genau zu
Ich kann Innovationen auch gegenüber skeptischen Kollegen durchsetzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mein Beruf ist interessant, weil ich täglich neu herausgefordert werde.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Für jedes Problem kann ich eine Lösung finden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bezweifle, dass ich die vor mir liegenden Aufgaben bewältigen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich weiß, dass ich zu den Eltern guten Kontakt halten kann, selbst in schwierigen Situationen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Schwierigkeiten sehe ich gelassen entgegen, weil ich meinen Fähigkeiten immer vertrauen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin niedergeschlagen, weil meine berufliche Situation so schlecht ist.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Auch bei überraschenden Ereignissen glaube ich, dass ich gut mit ihnen zurechtkommen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich befürchte, dass mich die beruflichen Belastungen überfordern.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Denken Sie bitte bei den folgenden Fragen an Situationen, die Sie im Referendariat aufregen oder belasten.

	trifft nicht zu					trifft genau zu
Ich versuche, die belastende Situation aktiv zu verändern.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich spreche mit Freunden und Bekannten über die Schwierigkeiten.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich versuche, das Problem alleine zu lösen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich neige dann schnell dazu, zu resignieren.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich suche mir KollegInnen an meiner Ausbildungsschule, die mir bei der Lösung der Probleme helfen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich versuche, die Probleme aktiv zu bewältigen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fahre so schnell wie möglich nach Hause, um mich dort von den Problemen in der Schule zu erholen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich versuche, die Probleme zu ignorieren.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich rede mit meinem Partner/meiner Partnerin darüber.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich suche mir andere ReferendarInnen, die mir bei der Lösung der Probleme helfen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bitte KollegInnen an meiner Ausbildungsschule um Hilfe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich versuche, die Ursachen der Belastungen zu beseitigen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich versuche, mir nichts anmerken zu lassen und „normal“ weiterzuarbeiten.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich versuche, der Situation auszuweichen und mich zu Hause oder in den Ferien zu erholen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	trifft nicht zu				trifft genau zu	
Ich bitte die AusbilderInnen an meinem Studienseminar um Hilfe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich ärgere mich über die Schulleitung und/oder die Kollegen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich hole mir Rat bei meinem/meiner MentorIn.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

In den folgenden Aussagen geht es um die Beurteilung Ihres Gesundheitszustandes. Bitte beantworten Sie jede der Fragen, indem Sie bei den Antwortmöglichkeiten den Kreis ankreuzen, der am besten auf Sie zutrifft.

	schlecht				ausgezeichnet	
Wie würden Sie Ihren Gesundheitszustand im allgemeinen beschreiben?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	derzeit schlechter				derzeit besser	
Im Vergleich zum vergangenen Monat, wie würden Sie Ihren derzeitigen Gesundheitszustand beschreiben?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Im Folgenden sind einige körperliche Beschwerden aufgeführt. Überlegen Sie bitte, wie häufig Sie unter den folgenden Beschwerden im vergangenen Monat litten.

	nie				sehr häufig	
Magenschmerzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Schwächegefühl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Kopfschmerzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Herzklopfen, Herzjagen oder Herzstolpern	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Schwindelgefühl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mattigkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sodbrennen oder saures Aufstoßen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nacken- und Schulterschmerzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Stimmprobleme	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Psychische Probleme	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	eher in der Schule		keine Unterschiede		eher in den Ferien	
Haben Sie den Eindruck, dass die o.g. Beschwerden eher während der Schulzeit oder in den Ferien auftreten?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

In diesen Fragen geht es darum, wie Sie sich fühlen und wie es Ihnen im vergangenen Monat gegangen ist.

	immer				nie	
Wie oft waren Sie im vergangenen Monat ...voller Schwung?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...sehr nervös?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... so niedergeschlagen, dass Sie nichts aufheitern konnte?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ruhig und gelassen?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Anhang

Wie oft waren Sie im vergangenen Monat	immer					nie
...voller Energie?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...entmutigt und traurig?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...erschöpft?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...glücklich?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...müde?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie häufig haben Ihre körperliche Gesundheit oder seelischen Probleme im vergangenen Monat Ihre Sozialkontakte (Besuche bei Freunden, Verwandten usw.) beeinträchtigt?	immer					nie
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Inwieweit trifft jede der folgenden Aussagen auf Sie zu?

	trifft nicht zu					trifft genau zu
Ich scheine etwas leichter als andere krank zu werden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin genauso gesund wie alle anderen, die ich kenne.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich übernehme häufig freiwillig zusätzliche Aufgaben, zu denen ich nicht verpflichtet bin.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn KollegInnen an meiner Ausbildungsschule überlastet sind, helfe ich, wo ich kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe wenig Lust, mich in der Schule besonders zu engagieren, weil man dafür ja doch keine Anerkennung bekommt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich lese regelmäßig pädagogische Fachzeitschriften.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich beteilige mich aktiv an Besprechungen und Konferenzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich informiere mich regelmäßig über neue Unterrichtskonzepte.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bei Fragen von KollegInnen meiner Ausbildungsschule oder Unklarheiten bin ich stets bereit zu helfen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bei der Organisation von Schulfesten und ähnlichen Veranstaltungen wirke ich tatkräftig mit.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Im Folgenden bitten wir Sie um die Beantwortung einiger Fragen zu Ihrer konkreten Situation und Ihrer persönlichen Einschätzung hinsichtlich Beruf und Leben.

	trifft nicht zu					trifft genau zu
Ich fühle mich durch alltägliche Hausarbeit (Einkaufen, Putzen etc.) belastet.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fühle mich durch Kindererziehung und -betreuung belastet.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Zum Feierabend ist die Arbeit für mich vergessen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Meine bisherige Ausbildung war recht erfolgreich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mit meinem bisherigen Leben kann ich zufrieden sein.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	trifft nicht zu					trifft genau zu
Mein Partner/meine Partnerin zeigt Verständnis für meine Arbeit.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Berufliche Karriere bedeutet mir wenig.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Im großen und ganzen bin ich glücklich und zufrieden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich könnte auch ohne meine Arbeit ganz glücklich sein.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Bei meiner Arbeit habe ich den Ehrgeiz, keinerlei Fehler zu machen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Misserfolge werfen mich nicht um, sondern veranlassen mich zu noch stärkerer Anstrengung.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich glaube, dass ich ziemlich hektisch bin.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wirkliche Erfolge in meiner Ausbildung sind mir bisher versagt geblieben.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Von meinem Partner/meiner Partnerin wünsche ich mir mehr Rücksichtnahme auf meine beruflichen Aufgaben und Probleme.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich strebe nach höheren beruflichen Zielen als die meisten anderen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich neige dazu, über meine Kräfte hinauszuarbeiten.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Was immer ich tue, es muss perfekt sein.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Feierabend ist Feierabend, da verschwende ich keinen Gedanken mehr an die Arbeit.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin mir sicher, dass ich auch die künftigen Anforderungen des Lebens gut bewältigen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich glaube, ich bin ein ruhender Pol in meinem Umfeld.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich wüsste nicht, wie ich ohne Arbeit leben sollte.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
In der Arbeit verausgabe ich mich stark.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich irgendwo versagt habe, kann mich das ziemlich mutlos machen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn mir etwas nicht gelingt, bleibe ich hartnäckig und strenge mich umso mehr an.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Berufliche Fehlschläge können mich leicht entmutigen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mit der Anleitung an meiner Ausbildungsschule bin ich zufrieden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mit der Ausbildung an meinem Studienseminar bin ich zufrieden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Tragen Sie nun noch einen Code ein. Dieser Code ermöglicht Ihnen, Ihre persönlichen Ergebnisse bei uns abzurufen. Der Code ist dergestalt formuliert, daß ein Rückschluss auf Ihre Person nicht möglich ist.

Erster Buchstabe des Vornamens der Mutter: z.B. **J** für **J**ohanna

Geburtsdatum der Mutter: z.B. **210538** für **21. Mai 1938**

Vielen Dank!

Wie groß ist in etwa das Kollegium an Ihrer Ausbildungsschule? _____

Schätzen Sie bitte den durchschnittlichen Ausländeranteil in den Klassen, in denen Sie unterrichten (in %): _____

An wie vielen Klassenfahrten haben Sie bislang während Ihres Referendariats teilgenommen? _____

An wie vielen Schultagen haben Sie in den letzten drei Monaten wegen Krankheit gefehlt? an _____ Tagen

Beteiligen Sie sich an der Gremienarbeit an Ihrem Studienseminar (Studienseminarrat/Personalrat)? nein ja

Wenn ja, wie groß schätzen Sie den Zeitaufwand dafür pro Woche ein? _____ Stunden/Woche

Führen Sie parallel zu dem Referendariat einen Nebenjob aus? nein ja

Wenn ja, wie viele Stunden arbeiten Sie im Rahmen dieses Nebenjobs pro Woche? _____ Stunden/Woche

Waren Sie vor dem Referendariat berufstätig (ausgenommen Studium)? nein ja

Wann haben Sie mit dem Referendariat begonnen (Monat/Jahr)? _____ / _____

Im Folgenden sehen Sie eine Liste von Arbeitsbedingungen, die im Schulalltag bzw. während des Referendariats auftreten. Bitte kreuzen Sie bei jeder Situation an, wie belastend diese Situation aktuell für Sie ist. Ein Kreuz ganz rechts würde bedeuten, dass Sie die geschilderte Situation aktuell als sehr belastend empfinden; ein Kreuz ganz links würde bedeuten, dass Sie die geschilderte Situation aktuell als überhaupt nicht belastend empfinden. Die anderen Kreise bedeuten jeweils, dass Sie die Situation als eher belastend oder eher nicht belastend empfinden.

	überhaupt nicht belastend					sehr belastend
Zu große Klassen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mangelnde Motivation bzw. Konzentrationsfähigkeit der SchülerInnen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Disziplinprobleme	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Probleme mit ausländischen SchülerInnen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ärger mit Behörden bzw. Institutionen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Probleme mit den Eltern der SchülerInnen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fachfremder Unterrichtseinsatz	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Unterschiedliche Lernvoraussetzungen der SchülerInnen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Vor- und Nachbereitung des Unterrichts	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Geringe Lernbereitschaft von SchülerInnen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verwaltungsarbeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ständige Kritik am Lehrerberuf und fehlende Anerkennung in der Öffentlichkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hektisches und störungsvolles Arbeitsklima an der Schule	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Unterrichtsbesuche / Lehrproben	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Konflikte mit der Ausbildungsschule	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	überhaupt nicht belastend				sehr belastend	
Eigenverantwortlicher Unterricht	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Konflikte mit AusbilderInnen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pädagogische Prüfungsarbeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Examensprüfungen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Klassenfahrten	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Seminarveranstaltungen am Studienseminar	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Einstellungsdruck	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Unterrichtsentwürfe	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Schlechte Besoldung des Referendariats	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	überhaupt nicht				sehr	
Wie sehr belasten Sie außerberufliche Anforderungen?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Bitte füllen Sie nun die folgende Tabelle aus. Bitte tragen Sie mit Zahlen von 1-6 ein, wie sehr die Aussagen für Sie zutreffen. Die 1 bedeutet „trifft gar nicht zu“, die 6 bedeutet „trifft genau zu“. Die Zahlen 2-5 entsprechen den jeweiligen Zwischenabstufungen.

Mit Karriere ist hier nicht nur beruflicher Aufstieg gemeint, da dies im Lehrberuf nur eingeschränkt möglich ist, sondern die individuelle Weiterqualifikation hinsichtlich pädagogischer und didaktischer Fertigkeiten. Mit Schule ist Ihre Ausbildungsschule und mit der Berufsgruppe der Lehrer die gesamte Gruppe – sowohl Lehrerinnen und Lehrer, als auch Referendarinnen und Referendare - gemeint.

Beispiel: Wenn Sie sich sehr mit Ihrer Karriere (im Sinne eigener beruflicher Weiterqualifikation) identifizieren, tragen Sie in das entsprechende Feld eine 6 ein, identifizieren Sie sich wenig oder gar nicht mit der Berufsgruppe der Lehrer, tragen Sie in die entsprechenden Felder eine 2 oder 1 ein usw.

Bitte füllen Sie jedes Feld der Tabelle aus.

<i>1 = trifft gar nicht zu</i> <i>6 = trifft genau zu</i>	Karriere	Ausbildungsschule	Berufsgruppe der Lehrer
Ich identifiziere mich mit...			
Karriereorientiert bzw. Angehöriger meiner Ausbildungsschule bzw. Berufsgruppe zu sein, spiegelt meine Persönlichkeit gut wider.			
Für die Karriere bzw. in meiner Ausbildungsschule/ meinem Beruf zu arbeiten, macht mir viel Spaß.			
Ich denke eher ungern an meine...			
Die Karriere, meine Schule/Berufsgruppe wird von anderen positiv bewertet.			
Für meine Karriere/Schule/Berufsgruppe engagiere ich mich auch über das absolut Notwendige hinaus.			

Wenn im Referendariat Probleme auftreten, kann es wichtig sein, ob man sich durch verschiedene Personen unterstützt fühlt. Diese Unterstützung kann zum einen ganz praktischer Art sein, z.B. in Form wertvoller Ratschläge. Es kann sich aber auch um emotionale Unterstützung handeln, wenn man z.B. merkt, dass einem Vertrauen und Anerkennung entgegen gebracht werden oder man sich einfach bei jemandem aussprechen kann. Bitte geben Sie in der untenstehenden Tabelle möglichst genau an, als **wie stark** Sie die **praktische** bzw. **emotionale** Unterstützung durch die genannten Gruppen / Personen empfinden. Dazu stehen Ihnen Angaben von 1 bis 6 zur Verfügung. Eine 1 bedeutet dabei, dass Sie von der entsprechenden Gruppe/Person "überhaupt nicht" unterstützt werden, eine 6 bedeutet, dass Sie "voll und ganz" unterstützt werden. Die Zwischenstufen bedeuten dementsprechend eher schwächere (2 und 3) oder eher stärkere (4 und 5) Unterstützung.

1 = überhaupt keine Unterstützung 6 = völlige Unterstützung	AusbilderInnen	MentorInnen/ BetreuerInnen	ReferendarskollegInnen	KollegInnen an der Ausbildungsschule	Partner oder Partnerin und Freunde	Schüler
Wie sehr werden Sie von folgenden Personen praktisch unterstützt, wenn die Arbeit im Zusammenhang mit dem Referendariat schwierig wird?						
Wie sehr fühlen Sie sich von folgenden Personen emotional unterstützt, wenn Probleme im Zusammenhang mit dem Referendariat auftauchen?						

Bitte geben Sie nun an, ob die folgenden Aussagen auf Sie persönlich zutreffen.

	trifft nicht zu					trifft genau zu
Ich bin zuversichtlich, dass ich den Anforderungen im Beruf gewachsen bin.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin mir sicher, dass ich mich in Zukunft auf individuelle Probleme der Schüler noch besser einstellen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn sich Widerstände auftun, finde ich Mittel und Wege, mich durchzusetzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin bedrückt, weil ich momentan einfach nicht klarkomme.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich traue mir zu, die Schüler für neue Projekte zu begeistern.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn eine neue Sache auf mich zukommt, weiß ich, wie ich damit umgehen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich mache mir Sorgen wegen der Probleme, die auf mich zukommen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin mir sicher, dass ich auch mit den problematischen Schülern in guten Kontakt kommen kann, wenn ich mich darum bemühe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
In letzter Zeit ist alles schief gelaufen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	trifft nicht zu					trifft genau zu
Ich kann Innovationen auch gegenüber skeptischen Kollegen durchsetzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mein Beruf ist interessant, weil ich täglich neu herausgefordert werde.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Für jedes Problem kann ich eine Lösung finden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bezweifle, dass ich die vor mir liegenden Aufgaben bewältigen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich weiß, dass ich zu den Eltern der SchülerInnen guten Kontakt halten kann, selbst in schwierigen Situationen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Schwierigkeiten sehe ich gelassen entgegen, weil ich meinen Fähigkeiten immer vertrauen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin niedergeschlagen, weil meine berufliche Situation so schlecht ist.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Auch bei überraschenden Ereignissen glaube ich, dass ich gut mit ihnen zurechtkommen kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich befürchte, dass mich die beruflichen Belastungen überfordern.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Denken Sie bitte bei den folgenden Fragen an Situationen, die Sie im allgemeinen oder im Zusammenhang mit dem Referendariat aufregen oder belasten.

	trifft nicht zu					trifft genau zu
Ich versuche, die belastende Situation aktiv zu verändern.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich spreche mit Freunden und Bekannten über die Schwierigkeiten.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich versuche, das Problem alleine zu lösen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich neige dann schnell dazu, zu resignieren.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich suche mir KollegInnen an meiner Ausbildungsschule, die mir bei der Lösung der Probleme helfen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich versuche, die Probleme aktiv zu bewältigen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fahre so schnell wie möglich nach Hause, um mich dort von den Problemen in der Schule zu erholen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich versuche, die Probleme zu ignorieren.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich rede mit meinem Partner/meiner Partnerin darüber.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich suche mir andere ReferendarInnen, die mir bei der Lösung der Probleme helfen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bitte KollegInnen an meiner Ausbildungsschule um Hilfe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich versuche, die Ursachen der Belastungen zu beseitigen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich versuche, mir nichts anmerken zu lassen und „normal“ weiterzuarbeiten.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich versuche, der Situation auszuweichen und mich zu Hause oder in den Ferien zu erholen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	trifft nicht zu				trifft genau zu	
Ich bitte die AusbilderInnen an meinem Studienseminar um Hilfe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich ärgere mich über die Schulleitung und/oder die Kollegen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich hole mir Rat bei meinem/meiner MentorIn.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

In den folgenden Aussagen geht es um die Beurteilung Ihres Gesundheitszustandes. Bitte beantworten Sie jede der Fragen, indem Sie bei den Antwortmöglichkeiten den Kreis ankreuzen, der am besten auf Sie zutrifft.

	schlecht				ausgezeichnet	
Wie würden Sie Ihren Gesundheitszustand im Allgemeinen beschreiben?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	derzeit schlechter				derzeit besser	
Im Vergleich zum vergangenen Monat, wie würden Sie Ihren derzeitigen Gesundheitszustand beschreiben?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Im folgenden sind einige körperliche Beschwerden aufgeführt. Überlegen Sie bitte, wie häufig Sie unter den folgenden Beschwerden im vergangenen Monat litten.

	nie				sehr häufig	
Magenschmerzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Schwächegefühl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Kopfschmerzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Herzklopfen, Herzjagen oder Herzs stolpern	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Schwindelgefühl	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mattigkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sodbrennen oder saures Aufstoßen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nacken- und Schulterschmerzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Stimmprobleme	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Psychische Probleme	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	eher in der Schule		keine Unterschiede		eher in den Ferien	
Haben Sie den Eindruck, dass die o.g. Beschwerden eher während der Schulzeit oder in den Ferien auftreten?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

In diesen Fragen geht es darum, wie Sie sich fühlen und wie es Ihnen im vergangenen Monat gegangen ist.

	nie				immer	
Wie oft waren Sie im vergangenen Monat ...voller Schwung?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...sehr nervös?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...so niedergeschlagen, dass Sie nichts aufheitern konnte?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...ruhig und gelassen?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Anhang

Wie oft waren Sie im vergangenen Monat	nie					immer
...voller Energie?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...entmutigt und traurig?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...erschöpft?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...glücklich?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...müde?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	nie					immer
Wie häufig haben Ihre körperliche Gesundheit oder seelischen Probleme im vergangenen Monat Ihre Sozialkontakte (Besuche bei Freunden, Verwandten usw.) beeinträchtigt?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Inwieweit trifft jede der folgenden Aussagen auf Sie zu?

	trifft nicht zu					trifft genau zu
Ich scheine etwas leichter als andere krank zu werden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin genauso gesund wie alle anderen, die ich kenne.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mit der Anleitung an meiner Ausbildungsschule bin ich zufrieden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mit der Ausbildung an meinem Studienseminar bin ich zufrieden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Tragen Sie nun noch einen Code ein. Dieser persönliche Code soll die Zuordnung dieses Fragebogens zu der letzten und den weiteren Befragungen ermöglichen. Wie bereits eingangs erwähnt, gewährleistet diese Methode Ihre Anonymität.

Bitte tragen Sie hier den Geburtstag Ihrer Mutter und Ihr eigenes Geburtsdatum ein (jeweils nur Tag und Monat).

Geburtsstag der Mutter: Tag _____ Monat _____

Eigener Geburtsstag: Tag _____ Monat _____

Vielen Dank für ihre Mitarbeit!

Wir werden Sie Mitte Mai erneut kontaktieren, um eine weitere Befragung zu diesem Thema durchzuführen.

Skalen mit Itembezeichnungen

Tabelle A1. Originalitems und Itembezeichnung der Neurotizismusskala

Item	Itembezeichnung
Ich bin leicht beunruhigt.	neuro01
Ich fühle mich anderen oft unterlegen.	neuro02
Wenn ich stark unter Stress stehe, fühle ich mich manchmal, als ob ich zusammenbräche.	neuro03
Ich fühle mich selten einsam oder traurig. (Contraire)	neuro04
Ich fühle mich oft angespannt und nervös.	neuro05
Manchmal fühle ich mich völlig wertlos.	neuro06
Ich empfinde selten Furcht oder Angst. (Contraire)	neuro07
Ich ärgere mich oft darüber, wie andere Leute mich behandeln.	neuro08
Zu häufig bin ich entmutigt und will aufgeben, wenn etwas schief geht.	neuro09
Ich bin selten traurig oder deprimiert. (Contraire)	neuro10
Ich fühle mich oft hilflos und wünsche mir eine Person, die meine Probleme löst.	neuro11
Manchmal war mir etwas so peinlich, dass ich mich am liebsten versteckt hätte.	neuro12

Tabelle A2. Originalitems und Itembezeichnung der Skala zu Messung des Ausmaßes an sozialer Unterstützung

1 = überhaupt keine Unterstützung 6 = völlige Unterstützung	Ausbil-derInnen	Mentor-Innen	Referendars-kollegInnen	KollegInnen an der Ausbil-dungsschule	Partner oder Partnerin und Freunde	Schüler
Wie sehr werden Sie von folgende Personen praktisch unterstützt, wenn die Arbeit im Zusammenhang mit dem Referendariat schwierig wird?	sozgr01	sozgr02	sozgr03	sozgr04	sozgr05	sozgr06
Wie sehr fühlen Sie sich von folgenden Personen emo-tional unterstützt, wenn Probleme im Zusammenhang mit dem Referendariat auftauchen?	sozgr07	sozgr08	sozgr09	sozgr10	sozgr11	sozgr12

Tabelle A3. Originalitems und Itembezeichnung der Skala zur Messung der Herausforderungs-, Bedrohungs- und Verlusteinschätzung und der Selbstwirksamkeitserwartungen (Reihenfolge wie im Fragebogen)

Originalitem	Itembezeichnung
Ich bin zuversichtlich, dass ich den Anforderungen im Beruf gewachsen bin.	chall01
Ich bin mir sicher, dass ich mich in Zukunft auf individuelle Probleme der Schüler noch besser einstellen kann.	wirkl01
Wenn sich Widerstände auftun, finde ich Mittel und Wege, mich durchzusetzen.	wirka01
Ich bin bedrückt, weil ich momentan einfach nicht klarkomme.	loss01
Ich traue mir zu, die Schüler für neue Projekte zu begeistern.	wirkl09
Wenn eine neue Sache auf mich zukommt, weiß ich, wie ich damit umgehen kann.	wirka09
Ich mache mir Sorgen wegen der Probleme, die auf mich zukommen.	threat01
Ich bin mir sicher, dass ich auch mit den problematischen Schülern in guten Kontakt kommen kann, wenn ich mich darum bemühe.	wirkl03
In letzter Zeit ist alles schief gelaufen.	loss03
Ich kann Innovationen auch gegenüber skeptischen Kollegen durchsetzen.	wirkl10
Mein Beruf ist interessant, weil ich täglich neu herausgefordert werde.	chall02
Für jedes Problem kann ich eine Lösung finden.	wirka08
Ich bezweifle, dass ich die vor mir liegenden Aufgaben bewältigen kann.	threat03
Ich weiß, dass ich zu den Eltern guten Kontakt halten kann, selbst in schwierigen Situationen.	wirkl02
Schwierigkeiten sehe ich gelassen entgegen, weil ich meinen Fähigkeiten immer vertrauen kann.	wirka06
Ich bin niedergeschlagen, weil meine berufliche Situation so schlecht ist.	loss02
Auch bei überraschenden Ereignissen glaube ich, dass ich gut mit ihnen zurechtkommen kann.	wirka05
Ich befürchte, dass mich die beruflichen Belastungen überfordern.	threat02

Anmerkungen: chall = Herausforderung; threat = Bedrohung; loss = Verlust, wirka = Allgemeine Selbstwirksamkeitserwartungen; wirkl = situationsspezifische Selbstwirksamkeitserwartungen

Tabelle A4. Originalitems, Itembezeichnung und Ladung der Items auf dem Copingfaktor

Originalitem	Itembezeichnung	Ladung auf dem Copingfaktor ¹
Ich versuche, die belastende Situation aktiv zu verändern.	cope01	.41
Ich spreche mit Freunden und Bekannten über die Schwierigkeiten.	cope02	.29
Ich versuche, das Problem alleine zu lösen.	cope04	-.18
Ich neige dann schnell dazu, zu resignieren.	<i>cope08</i>	-.37
Ich suche mir KollegInnen an meiner Ausbildungsschule, die mir bei der Lösung der Probleme helfen.	cope10	.70
Ich versuche, die Probleme aktiv zu bewältigen.	cope11	.52
Ich fahre so schnell wie möglich nach Hause, um mich dort von den Problemen in der Schule zu erholen.	<i>cope13</i>	-.33
Ich versuche, die Probleme zu ignorieren.	<i>cope14</i>	-.37
Ich rede mit meinem Partner/meiner Partnerin darüber.	cope16	.15
Ich suche mir andere ReferendarInnen, die mir bei der Lösung der Probleme helfen.	copen01	.72
Ich bitte KollegInnen an meiner Ausbildungsschule um Hilfe.	cope23	.50
Ich versuche, die Ursachen der Belastungen zu beseitigen.	<i>cope24</i>	-.36
Ich versuche, mir nichts anmerken zu lassen und „normal“ weiterzuarbeiten.	<i>cope26</i>	-.50
Ich versuche, der Situation auszuweichen und mich zu Hause oder in den Ferien zu erholen.	cope27	-.25
Ich bitte die AusbilderInnen an meinem Studienseminar um Hilfe.	copen02	.34
Ich ärgere mich über die Schulleitung und/oder die Kollegen.	cope28	.19
Ich hole mir Rat bei meinem/meiner MentorIn.	copen03	-. ²

Anmerkung: kursive Itembezeichnung = Items wurden für alle weiteren Analysen rekodiert; 1 = Faktorladung vor Rekodierung der Items; 2 = Item wurde aus allen Analysen ausgeschlossen (s. S. 84) fett gedruckte Faktorladungen = Items wurden für alle weiteren Analysen verwendet; copen = Eigenentwicklung;

Tabelle A5. Originalitems und Itembezeichnung der Beschwerdeskala

Originalitem	Itembezeichnung
Magenschmerzen	beschw01
Schwächegefühl	beschw02
Kopfschmerzen	beschw03
Herzklopfen, Herzjagen oder Herzstolpern	beschw04
Schwindelgefühl	beschw05
Mattigkeit	beschw06
Sodbrennen oder saures Aufstoßen	beschw07
Nacken- und Schulterschmerzen	beschw08
Stimmprobleme ¹	beschw09

Anmerkung: 1 = Eigenentwicklung

Tabelle A6. Originalitems und Itembezeichnung der Skala zur Messung des psychischen Wohlbefindens

	Originalitem	Itembezeichnung
Wie oft waren Sie im vergangenen Monat...	...voller Schwung?	psych01
	...sehr nervös? (Contrait)	psych02
	... so niedergeschlagen, dass Sie nichts aufheitern konnte? (Contrait)	psych03
	...ruhig und gelassen?	psych04
	...voller Energie?	psych05
	...entmutigt und traurig? (Contrait)	psych06
	...erschöpft? (Contrait)	psych07
	...glücklich?	psych08
	...müde? (Contrait)	psych09

Zusätzliche Dokumentation von Ergebnissen aus Studie 1

Tabelle A7. Rücklauf für alle Studienseminare in Studie 1

Studienseminar	Potenzial	Rücklauf	Rücklaufquote (%)
1	96	39	40.63
2	75	52	69.33
3	107	47	43.93
4	94	47	50.00
5	95	34	35.79
6	71	58	81.69
7	98	47	47.96
8	80	43	53.75
9	78	22	28.21
10	107	44	41.12
11	94	20	21.28
12	91	19	20.88
13	81	21	25.93
14	77	64	83.12
15	100	63	63.00
16	87	10	11.49
17	83	14	16.87
<i>Gesamt</i>	<i>1514</i>	<i>644</i>	<i>42.54</i>

Tabelle A8. Item-zu-Parcel-Zuordnung in Studie 1 und Studie 2

Skala	Parcels			
	1	2	3	4
Neurotizismus	neuro01, neuro07, neuro11, neuro12	neuro01, neuro06, neuro08, neuro09	neuro03, neuro04, neuro05, neuro10	-
Ausmaß an sozialer Unterstützung	sozgr01, sozgr03, sozgr05	sozgr06, sozgr04	sozgr10, sozgr09, sozgr11	sozgr07, sozgr12
Ereigniseinschätzung	threat02, loss02, chall02r	chall01r, threat01, loss03	threat03, loss01	
Selbstwirksamkeits- erwartungen	wirkl09	wirkl10, wirkl02	wirkl03, wirkl01	
Coping	cope11, cope01, cope13 ¹	cope24, cope23, cope14 ¹	cope10, cope27 ¹ , cope08 ¹ , cope26 ¹	
Körperliche Beschwerden	beschw05, beschw02, beshcw04	beschw06, beschw08	beschw03, beschw01	
Psychisches Wohlbefinden	psych03, psych02	psych04, psych08	psych01, psych05, psych06	

Anmerkung: 1 = Items wurden vor der Erstellung der Parcel rekodiert.

Zusätzliche Dokumentation von Ergebnissen aus Studie 2

Tabelle A9. Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t1 für den Längsschnittdatensatz L1

Parcel	Faktorladungen		
	Neurotizismus	Beschwerden	Psychisches Wohlbefinden
1	.81	.59	.66
2	.74	.75	.66
3	.75	.69	.66
Interkorrelationen			
	1	2	3
1 Neurotizismus	1	.58***	-.50***
2 Beschwerden		1	-.67***
3 Psychisches Wohlbefinden			1

Anmerkung: *** $p < .001$

Tabelle A10. Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t2 für den Längsschnittdatensatz L1

Parcel	Faktorladungen					
	Soziale Unterstützung	Ereigniseinschätzung	Selbstwirksamkeitseinschätzung	Coping	Beschwerden	Psychisches Wohlbefinden
1	.80	.80	.73	.79	.84	.71
2	.84	.89	.64	.71	.64	.79
3	.73	.84	.73	.76	.60	.91
4	.58	-	-	-	-	-

	Interkorrelationen					
	1	2	3	4	5	6
1 Soziale Unterstützung	1	-.33*	.48**	.50***	-.06	.30*
2 Ereigniseinschätzung		1	-.72***	-.59***	.38**	-.80***
3 Selbstwirksamkeitserwartungen			1	.56***	-.07	.59***
4 Coping				1	-.26	.52***
5 Beschwerden					1	-.54***
6 Psychisches Wohlbefinden						1

Anmerkungen: * p < .05; ** p < .01; *** p < .001

Tabelle A11. Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t3 für den Längsschnittdatensatz L1

Parcel	Faktorladungen					
	Soziale Unterstützung	Ereigniseinschätzung	Selbstwirksamkeitseinschätzung	Coping	Beschwerden	Psychisches Wohlbefinden
1	.62	.85	.53	.79	.77	.74
2	.82	.90	.63	.78	.79	.82
3	.67	.84	.68	.80	.78	.83
4	.71	-	-	-	-	-

	Interkorrelationen					
	1	2	3	4	5	6
1 Soziale Unterstützung	1	-.48**	.56**	.59**	-.04	.43**
2 Ereigniseinschätzung		1	-.81***	-.74***	.29*	-.82***
3 Selbstwirksamkeitserwartungen			1	.70***	-.15	.67***
4 Coping				1	-.27*	.64***
5 Beschwerden					1	-.52***
6 Psychisches Wohlbefinden						1

Anmerkungen: * p < .05; ** p < .01; *** p < .001

Tabelle A12. Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t4 für den Längsschnittdatensatz L1

Parcel	Faktorladungen					
	Soziale Unterstützung	Ereigniseinschätzung	Selbstwirksamkeitseinschätzung	Coping	Beschwerden	Psychisches Wohlbefinden
1	.73	.81	.80	.71	.74	.81
2	.77	.93	.65	.81	.77	.80
3	.79	.81	.80	.79	.62	.93
4	.64	-	-	-	-	-

	Interkorrelationen					
	1	2	3	4	5	6
1 Soziale Unterstützung	1	-.40	.47	.47	-.11	.37**
2 Ereigniseinschätzung		1	-.68***	-.68***	.59***	-.80***
3 Selbstwirksamkeitserwartungen			1	.47***	-.13	.50***
4 Coping				1	-.32*	.54***
5 Beschwerden					1	-.62***
6 Psychisches Wohlbefinden						1

Anmerkungen: * p < .05; ** p < .01; *** p < .001

Tabelle A13. Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t1 für den Längsschnittdatensatz L2

Parcel	Faktorladungen		
	Neurotizismus	Beschwerden	Psychisches Wohlbefinden
1	.82	.68	.74
2	.82	.70	.76
3	.81	.50	.75

	Interkorrelationen		
	1	2	3
1 Neurotizismus	1	.70***	-.70***
2 Beschwerden		1	-.66***
3 Psychisches Wohlbefinden			1

Anmerkung: *** $p < .001$

Tabelle A14. Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t2 für den Längsschnittdatensatz L2

Parcel	Faktorladungen					
	Soziale Unterstützung	Ereigniseinschätzung	Selbstwirksamkeitseinschätzung	Coping	Beschwerden	Psychisches Wohlbefinden
1	.78	.85	.66	.75	.87	.77
2	.79	.88	.64	.77	.71	.83
3	.75	.86	.77	.77	.63	.87
4	.58	-	-	-	-	-

	Interkorrelationen					
	1	2	3	4	5	6
1 Soziale Unterstützung	1	-.35***	.41***	.49***	-.10	.29**
2 Ereigniseinschätzung		1	-.81***	-.59***	.55***	-.81***
3 Selbstwirksamkeitserwartungen			1		-.28**	.73***
4 Coping				1	-.32**	.61***
5 Beschwerden					1	-.58***
6 Psychisches Wohlbefinden						1

Anmerkungen: ** p < .01; *** p < .001

Tabelle A15. Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t3 für den Längsschnittdatensatz L2

Parcel	Faktorladungen					
	Soziale Unterstützung	Ereigniseinschätzung	Selbstwirksamkeitseinschätzung	Coping	Beschwerden	Psychisches Wohlbefinden
1	.67	.87	.67	.78	.78	.76
2	.79	.90	.68	.80	.74	.84
3	.70	.87	.74	.76	.64	.84
4	.74	-	-	-	-	-

	Interkorrelationen					
	1	2	3	4	5	6
1 Soziale Unterstützung	1	-.40***	.46***	.58***	-.00	.34***
2 Ereigniseinschätzung		1	-.80***	-.73***	.42***	-.82***
3 Selbstwirksamkeitserwartungen			1	.64***	-.23*	.67***
4 Coping				1	-.33**	.67***
5 Beschwerden					1	-.62***
6 Psychisches Wohlbefinden						1

Anmerkungen: * p < .05; ** p < .01; *** p < .001

Tabelle A16. Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t4 für den Längsschnittdatensatz L2

Parcel	Faktorladungen					
	Soziale Unterstützung	Ereigniseinschätzung	Selbstwirksamkeitseinschätzung	Coping	Beschwerden	Psychisches Wohlbefinden
1	.71	.83	.84	.75	.78	.84
2	.73	.93	.65	.81	.75	.81
3	.79	.87	.69	.83	.65	.88
4	.67	-	-	-	-	-

	Interkorrelationen					
	1	2	3	4	5	6
1 Soziale Unterstützung	1	-.41***	.52***	.58***	-.08	.40***
2 Ereigniseinschätzung		1	-.61***	-.76***	.59***	-.82***
3 Selbstwirksamkeitserwartungen			1	.53***	-.10	.49***
4 Coping				1	-.45***	.64***
5 Beschwerden					1	-.64***
6 Psychisches Wohlbefinden						1

Anmerkung: *** $p < .001$

Tabelle A17. Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t1 für den Längsschnittdatensatz L3

Parcel	Faktorladungen		
	Neurotizismus	Beschwerden	Psychisches Wohlbefinden
1	.84	.71	.79
2	.82	.74	.77
3	.80	.55	.74

	Interkorrelationen		
	1	2	3
1 Neurotizismus	1	.67***	-.68***
2 Beschwerden		1	-.65***
3 Psychisches Wohlbefinden			1

Anmerkungen: *** $p < .001$

Tabelle A18. Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t2 für den Längsschnittdatensatz L3

Parcel	Faktorladungen					
	Soziale Unterstützung	Ereigniseinschätzung	Selbstwirksamkeitseinschätzung	Coping	Beschwerden	Psychisches Wohlbefinden
1	.72	.82	.65	.74	.82	.78
2	.84	.90	.64	.78	.72	.80
3	.68	.86	.70	.81	.63	.85
4	.61	-	-	-	-	-

	Interkorrelationen					
	1	2	3	4	5	6
1 Soziale Unterstützung	1	-.30***	.44***	.49***	-.17*	.29***
2 Ereigniseinschätzung		1	-.76***	-.61***	.51***	-.80***
3 Selbstwirksamkeitserwartungen			1	.64***	-.29**	.71***
4 Coping				1	-.44***	.67***
5 Beschwerden					1	-.63***
6 Psychisches Wohlbefinden						1

Anmerkungen: * p < .05; ** p < .01; *** p < .001

Tabelle A19. Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t3 für den Längsschnittdatensatz L3

Parcel	Faktorladungen					
	Soziale Unterstützung	Ereigniseinschätzung	Selbstwirksamkeitseinschätzung	Coping	Beschwerden	Psychisches Wohlbefinden
1	.69	.88	.68	.80	.73	.77
2	.75	.89	.67	.78	.74	.84
3	.69	.86	.71	.75	.70	.85
4	.73	-	-	-	-	-

	Interkorrelationen					
	1	2	3	4	5	6
1 Soziale Unterstützung	1	-.36***	.41***	.53***	-.04	.26**
2 Ereigniseinschätzung		1	-.76***	-.74***	.38***	-.82***
3 Selbstwirksamkeitserwartungen			1	.66***	-.19	.58***
4 Coping				1	-.31***	.64***
5 Beschwerden					1	-.58***
6 Psychisches Wohlbefinden						1

Anmerkungen: ** p < .01; *** p < .001

Tabelle A20. Faktorladungen und Interkorrelation der Skalen zu Messzeitpunkt t4 für den Längsschnittdatensatz L3

Parcel	Faktorladungen					
	Soziale Unterstützung	Ereigniseinschätzung	Selbstwirksamkeitseinschätzung	Coping	Beschwerden	Psychisches Wohlbefinden
1	.70	.83	.82	.76	.77	.83
2	.72	.91	.65	.82	.76	.81
3	.78	.87	.67	.84	.62	.87
4	.65	-	-	-	-	-

	Interkorrelationen					
	1	2	3	4	5	6
1 Soziale Unterstützung	1	-.43***	.53***	.57***	-.07	.41***
2 Ereigniseinschätzung		1	-.63***	-.73***	.57***	-.78***
3 Selbstwirksamkeitserwartungen			1	.56***	-.11	.53***
4 Coping				1	-.41***	.64***
5 Beschwerden					1	-.64***
6 Psychisches Wohlbefinden						1

Anmerkung: *** p < .001

Anhang

Interkorrelation der Skalen über alle Messzeitpunkte für den Längsschnittdatensatz L2¹

	1	2	3	4	5	6	7	8	9	10	11
1 neuro1	1	.51***	.59***	-.18*	.60***	-.49***	-.46***	.49***	-.60***	-.17*	.51***
2 beschw1		1	-.48***	-.05	.37***	-.28***	-.37***	.62***	-.41***	-.04	.26***
3 psych1			1	.27***	-.47***	.50***	.41***	-.35***	.56***	.21**	-.35***
4 sozgr2				1	-.32***	.37***	.42***	-.09	.26***	.71***	-.25***
5 bewert2					1	-.65***	-.53***	.41***	-.72***	-.30***	.74***
6 wirk2						1	.50***	-.20**	.57***	.33***	-.58***
7 cope2							1	-.33***	.55***	.34***	-.49***
8 beschw2								1	-.47***	-.04	.22**
9 psych2									1	.23**	-.56***
10 sozgr3										1	-.35***
11 bewert3											1
12 wirk3											
13 cope3											
14 beschw3											
15 psych3											
16 sozgr4											
17 bewert4											
18 wirk4											
19 cope4											
20 beschw4											
21 psych4											

Anmerkungen: 1 Die Korrelationen wurden in SPSS 11.5 berechnet und basieren auf, mit der EM-Methode ersetzten, fehlenden Werten im Datensatz L2;

* p < .05; ** p < .01; *** p < .001

Fortsetzung auf der nächsten Seite

Fortsetzung Tabelle A21

	12	13	14	15	16	17	18	19	20	21
1 neuro1	-.43***	-.40***	.49***	-.53***	-.18*	.54***	-.45***	-.44***	.46***	-.47***
2 beschw1	-.26***	-.22**	.67***	-.35***	-.00	.34***	-.14	-.25***	.66***	-.36***
3 psych1	.40***	.30***	-.33***	.47***	.26***	-.47***	.44***	.39***	-.30***	.47***
4 sozgr2	.35***	.46***	.02	.26***	.64***	-.25***	.35***	.29***	-.09	.25***
5 bewert2	-.65***	-.52***	.38***	-.61***	-.28***	.71***	-.51***	-.55***	.41***	-.44***
6 wirk2	.80***	.43***	-.15*	.50***	.30***	-.55***	.80***	.47***	-.20**	-.44***
7 cope2	.46***	.73***	-.32***	.47***	.38***	-.51***	.44***	.72***	-.33***	.39***
8 beschw2	-.16*	-.27***	.75***	-.38***	.12	.30***	-.05	.27***	.74***	-.39***
9 psych2	.55***	.45***	-.35***	.63***	.22**	-.55***	.48***	.57***	-.41***	.54***
10 sozgr3	.36***	.50***	-.01	.30***	.74***	-.28***	.30***	.40***	-.10	.29***
11 bewert3	-.66***	-.65***	.34***	-.73***	-.26***	.80***	-.49***	-.60***	.39***	-.57***
12 wirk3	1	.51***	-.18*	.55***	.28***	-.56***	.71***	.52***	-.28***	.47***
13 cope3		1	-.27***	.59***	.48***	-.56***	.37***	.82***	-.29***	.45***
14 beschw3			1	-.50***	.04	.41***	-.04	-.28***	.81***	-.49***
15 psych3				1	.29***	-.71***	.47***	.60***	-.45***	.71***
16 sozgr4					1	-.32***	.37***	.48***	-.06	.33***
17 bewert4						1	-.57***	-.70***	.53***	-.75***
18 wirk4							1	.48***	-.15*	.47***
19 cope4								1	-.40***	.60***
20 beschw4									1	-.58***
21 psych4										1

Anmerkungen: * $p < .05$; ** $p < .01$; *** $p < .001$

Tabelle A21. Parameterschätzer jeweils für das einfaktorielles LGC-Modell für das Ausmaß an sozialer Unterstützung, Ereigniseinschätzung und Selbstwirksamkeitserwartungen für den Längsschnittdatensatz L2

Parameterschätzer		Ausmaß an sozialer Unterstützung	Ereigniseinschätzung	Selbstwirksamkeitserwartungen
Faktorladungen der Skalenwerte (auf dem latenten Intercept-Faktor)	Skala zu t2	.78	.84	.90
	Skala zu t3	.85	.86	.87
	Skala zu t4	.79	.79	.80
Intercept-Faktor	Mittelwert	3.87***	2.25***	4.20***
Varianz	Intercept-Faktor	.42***	.63***	.42***
	Fehler Skala t2	.26***	.26***	.10***
	Fehler Skala 3	.17***	.22***	.14***
	Fehler Skala 4	.25***	.38***	.23***

Anmerkung: *** $p < .001$

Tabelle A22. Parameterschätzer jeweils für das zweifaktorielle LGC-Modell für Coping, körperliche Beschwerden und psychisches Wohlbefinden für den Längsschnittdatensatz L2

Parameterschätzer		Coping	Ausmaß an körperlichen Beschwerden	Psychisches Wohlbefinden
Faktorladungen der Skalenwerte (auf dem latenten Intercept-Faktor)	Skala t1	-	.87	.77
	Skala t2	.86	.72	.65
	Skala t3	.81	.75	.66
	Skala t4	.78	.72	.55
Faktorladungen der Skalenwerte (auf dem latenten Slope-Faktor)	Skala t1	-	.00	.00
	Skala t2	.00	.34	.37
	Skala t3	.54	.30	.43
	Skala t4	.26	.60	.53

Fortsetzung auf der nächsten Seite

Fortsetzung Tabelle A23

Parameterschätzer		Coping	Ausmaß an körperlichen Beschwerden	Psychisches Wohlbefinden
Intercept-Faktor	Mittelwert	4.60***	2.37***	4.67***
Slope-Faktor	Mittelwert	-.13**	.50***	-.38***
Varianz	Intercept-Faktor	.37***	.49***	.28***
	Slope-Faktor	.17*	.35***	.09**
	Fehler Skala t1	-	.16***	.19***
	Fehler Skala t2	.13*	.34***	.29***
	Fehler Skala t3	.03	.30***	.24***
	Fehler Skala t4	.20*	.12	.38***

Anmerkungen: ** p < .01; *** p < .001

Tabelle A23. Interkorrelationen (standardisierte Parameter) der Skalen jeweils innerhalb eines Messzeitpunkts für das multivariate ALT-Modell für den Längsschnittdatensatz L2

	1	2	3	4	5	6	7	8	9	10	11
1 beschw1	1	-.38***	-	-	-	-	-	-	-	-	-
2 psych1		1	-	-.46***	-	-	-	-	-	-	-
3 sozgr2			1	-	-	-	-	-	-	-	-
4 bewert2				1	-	-	-	-	-	-	-
5 wirk2					1	-	-	-	-	-	-
6 cope2						1	-	-	-	-	-
7 beschw2							1	-.27***	-	-	-
8 psych2								1	-	-	-
9 sozgr3									1	-.19*	-
10 bewert3										1	-.29**
11 wirk3											1
12 cope3											
13 beschw3											
14 psych3											
15 sozgr4											
16 bewert4											
17 wirk4											
18 cope4											
19 beschw4											
20 psych4											

Fortsetzung auf der nächsten Seite

Anhang

Fortsetzung Tabelle A21

	12	13	14	15	16	17	18	19	20
1 beschw1	-	-	-	-	-	-	-	-	-
2 psych1	-	-	-	-	-	-	-	-	-
3 sozgr2	-	-	-	-	-	-	-	-	-
4 bewert2	-	-	-	-	-	-	-	-	-
5 wirk2	-	-	-	-	-	-	-	-	-
6 cope2	-	-	-	-	-	-	-	-	-
7 beschw2	-	-	-	-	-	-	-	-	-
8 psych2	-	-	-	-	-	-	-	-	-
9 sozgr3	.46**	-	-	-	-	-	-	-	-
10 bewert3	-.57***	-	-.46***	-	-	-	-	-	-
11 wirk3	.38**	-	-	-	-	-	-	-	-
12 cope3	1	-	.36**	-	-	-	-	-	-
13 beschw3		1	-.29***	-	-	-	-	-	-
14 psych3			1	-	-	-	-	-	-
15 sozgr4				1	-.22*	.28*	-	-	.19 ⁺
16 bewert4					1	-.17*	-.52***	.71***	-.67***
17 wirk4						1	-	-.65***	-
18 cope4							1	-	.36***
19 beschw4								1	-.35***
20 psych4									1

Anmerkungen: * p < .05; ** p < .01; *** p < .001