

4 INSTRUMENTARIUM

In diesem Teil der Arbeit werden alle Instrumente kurz beschrieben sowie ihre internen Konsistenzen und Retest–Reliabilitäten vorgestellt. Der vollständige Wortlaut aller Skalen, Item– und Skalenkennwerte sowie weitere Informationen finden sich im Anhang A.

Im Rahmen des Modellversuchs *Verbund Selbstwirksame Schulen* wurden mehrere psychometrische Meßinstrumente entwickelt, die in ihrer endgültigen Fassung erstmals in der Skalendokumentation zum Modellversuch (Schwarzer & Jerusalem, 1999) erschienen sind. Es handelt sich hierbei um die Skalen zur Lehrer–Selbstwirksamkeitserwartung und zur Kollektiven Selbstwirksamkeitserwartung. Ausführliche Beschreibungen dieser Meßinstrumente finden sich im Theorieteil der vorliegenden Arbeit.

Die psychometrischen Kennwerte aller im Modellversuch verwendeten Items bzw. Skalen finden sich ebenfalls im Anhang A.

Die Items der verschiedenen Skalen wurden im Fragebogen durchmischt vorgegeben. Die Summenwerte der Skalen wurden nach der Regel gebildet, daß ein Teilnehmer mindestens 75% der Items einer Skala beantwortet haben mußte, um einen Skalensummenwert zu erhalten.

4.1 Allgemeine Selbstwirksamkeitserwartung

Das Konzept der Allgemeinen Selbstwirksamkeitserwartung fragt nach der persönlichen Einschätzung eigener Kompetenzen, *allgemein* mit Schwierigkeiten und Barrieren im täglichen Leben zurechtzukommen.

Die 10–Item–Skala wurde von Matthias Jerusalem und Ralf Schwarzer entwickelt und seit 1981 in einer großen Anzahl von Studien im In– und Ausland erfolgreich eingesetzt. Es liegen Daten aus 14 verschiedenen Kulturen vor. Die Skala wurde bisher in 26 Sprachen übersetzt. Das Instrument hat sich als sparsam, reliabel und valide erwiesen. Reliabilitäten lagen zwischen Cronbachs $\alpha = .75$ und $.90$.

Im Rahmen des Modellversuchs wurde die Skala überarbeitet und besteht jetzt in der revidierten Fassung (Jerusalem & Schwarzer, 1999). Ein Beispielitem lautet: „Schwierigkeiten sehe ich gelassen entgegen, weil ich meinen Fähigkeiten immer vertrauen kann“.

Das Antwortformat ist vierstufig: stimmt nicht (1) – stimmt kaum (2) – stimmt eher (3) – stimmt genau (4).

Die Reliabilitäten lagen zu den drei Meßzeitpunkten bei $\alpha = .86$, $.86$ und $.85$. Die Restest–Reliabilitäten lagen bei $r_{12} = .74$, $r_{23} = .78$ und $r_{13} = .64$.

4.2 Lehrer–Selbstwirksamkeitserwartung

Um sinnvolle Aussagen über einen spezifischen Bereich machen zu können, sollten Fragen nach der Selbstwirksamkeitserwartung inhaltliche Spezifität aufweisen. Wenn man beispielsweise den Einfluß von Selbstwirksamkeitserwartungen im Lehrerberuf ergründen möchte, sollte man zunächst innerhalb des Berufsfelds Bereiche mit unterschiedlichen Kompetenzanforderungen ausfindig machen. Für den Lehrerberuf handelt es sich dabei vor allem um die Bereiche berufliche Leistung, berufliche Weiterentwicklung, soziale Interaktionen mit Schülern, Eltern und Kollegen sowie Umgang mit Berufsstress (vgl. Rudow, 1994). All diese Bereiche sollten in einem (bereichsspezifischen) Meßinstrument zur Lehrer–Selbstwirksamkeit angemessen vertreten sein.

Bei der Entwicklung der vorliegenden Skala wurden zunächst jeweils fünf bis neun Items für jeden der genannten Bereiche entwickelt und mehrfach revidiert. Alle Items weisen die eindeutige semantische Struktur auf, die für Selbstwirksamkeitserwartungen typisch ist. Sie enthalten die subjektive Gewißheit, etwas tun zu können, auch wenn eine Barriere im Weg steht. Eine 6–Item–Skala erfaßte den Bereich der allgemeinen *beruflichen Leistung* mit Formulierungen wie „Ich weiß, daß ich es schaffe, selbst den problematischsten Schülern den prüfungsrelevanten Stoff zu vermitteln“. Sieben weitere Aussagen zielten auf den Bereich *berufsbezogener sozialer Interaktionen* nach der Art „Ich weiß, daß ich zu den Eltern guten Kontakt halten kann, selbst in schwierigen Situationen“. Eine 9–Item–Skala wurde entwickelt, um auch den Bereich des *Umgangs mit Stress und Emotionen* zu erfassen. Ein Beispiel lautet: „Selbst wenn es mir mal nicht so gut geht, kann ich doch im Unterricht immer noch gut auf die Schüler eingehen“. Schließlich wurde eine Skala mit fünf Aussagen konstruiert, bei der es um die spezifische Selbstwirksamkeit zum *innovativen Handeln* gehen sollte. Eine Formulierung dafür ist: „Ich traue mir zu, die Schüler für neue Projekte zu begeistern“.

Diese 27 Items, die von einer Projektgruppe im Modellversuch *Verbund Selbstwirksamer Schulen* entworfen und dann empirisch erprobt wurden, dienten als primärer Itempool zur Konstruktion einer kurzen, ökonomischen Lehrer–Selbstwirksamkeits–Skala. Es war das Ziel, ein sparsames Instrument von ungefähr zehn Aussagen zu konstruieren, das generell bei Lehrern eingesetzt werden kann (Schwarzer & Schmitz, 1999; vgl. die gesamte Skala im Anhang A). Die Auswahl der Items erfolgte vor allem im Diskursverfahren aufgrund theoretischer Überlegungen zur Inhaltsvalidität, nur sekundär aufgrund der empirischen Itemkenn-

werte. Es war nicht das Ziel, eine stark homogene Skala zu entwickeln und somit die interne Konsistenz zu maximieren, sondern es erschien sinnvoll, die Validitätsoptimierung in den Vordergrund zu stellen, indem relativ heterogene Handlungsbereiche im Berufsleben von Lehrern angesprochen werden.

Das Antwortformat ist vierstufig: stimmt nicht (1) – stimmt kaum (2) – stimmt eher (3) – stimmt genau (4).

Die Reliabilitäten lagen zu den drei Meßzeitpunkten bei $\alpha = .81$, $.76$ und $.78$. Die Restest-Reliabilitäten lagen bei $r_{12} = .67$, $r_{23} = .76$ und $r_{13} = .65$.

4.3 Kollektive Selbstwirksamkeitserwartung

Bei der Kollektiven Selbstwirksamkeitserwartung geht es darum, überindividuelle Überzeugungen von der Handlungskompetenz einer Gruppe zu konzeptualisieren. So wie der einzelne optimistische Selbstüberzeugungen haben kann, soll dies auch für Gruppen gelten. Es wird zum Beispiel angenommen, daß ein Lehrerkollegium Vertrauen in die Kapazitätsreserven des Teams haben kann und somit auch eine optimistische Auffassung von der Bewältigung zukünftiger streßreicher Ereignisse, die die ganze Gruppe treffen.

Generell könnte man für die Erfassung dieses Konstrukts verschiedene Wege gehen, zum Beispiel den des Interviews mit Gruppen oder den der Beobachtung von Gruppeninteraktionen. Im Rahmen der vorliegenden Feldstudie wurde jedoch entschieden, das Individuum weiterhin als Erhebungseinheit zu betrachten und somit die *individuelle* Wahrnehmung der Bewältigungskompetenz einer Bezugsgruppe zu erfassen, um die Erhebungsmethode für alle Konstrukte einheitlich zu gestalten. Ein erster Entwurf bestand zunächst aus 24 Items, die zum zweiten Meßzeitpunkt erstmalig eingesetzt wurden. Zum dritten Meßzeitpunkt wurde die Skala aufgrund psychometrischer Analysen halbiert.

Bisher hat es zu diesem Konstrukt noch keine deutschsprachige Skala gegeben. Aus einem Pool von 24 Items wurde eine 12-Item-Skala entwickelt (Schwarzer & Jerusalem, 1999) mit Aussagen wie „Da wir dieselben pädagogischen Absichten verfolgen, können wir Lehrer auch mit ‚schwierigen‘ Schülern an dieser Schule klarkommen“.

Das Antwortformat ist vierstufig: stimmt nicht (1) – stimmt kaum (2) – stimmt eher (3) – stimmt genau (4).

Der Itempool wurde ab der zweiten Datenerhebungswelle vorgegeben. Die Reliabilitäten lagen zu den letzten beiden Meßzeitpunkten bei $\alpha = .91$ und $.92$. Die Restest-Reliabilität lag bei $r_{23} = .77$.

4.4 Maslach Burnout Inventar *MBI*

Das Maslach Burnout Inventar *MBI* wurde von Christina Maslach und Susan Jackson 1986 entwickelt. Es ist momentan das Instrument zur Messung von Burnout mit der weitesten Verbreitung (Maslach, 1995). Burnout tritt vor allem in sozialen Berufen auf, in denen der Umgang mit Klienten (Patienten, Schülern etc.) einen großen Bestandteil der täglichen Arbeit ausmacht.

Das *MBI* enthält drei Subskalen, die die drei Komponenten des Burnout-Syndroms erfassen sollen. Es handelt sich dabei um Emotionale Erschöpfung, Depersonalisation und Leistungsverlust im beruflichen Zusammenhang. Das Instrument wurde in der Übersetzung von Dirk Enzmann und Dieter Kleiber (1989) eingesetzt.

Unter *emotionaler Erschöpfung* versteht man einen Zustand, in dem Menschen sich so ausgelaugt fühlen, daß sie keine Begeisterung, Interesse oder Enthusiasmus in ihrem Berufsleben mehr empfinden können. Ein Beispielim lautet: „Durch meine Arbeit bin ich gefühlsmäßig am Ende“.

Die Reliabilitäten lagen zu den drei Meßzeitpunkten bei $\alpha = .88$, $.86$ und $.87$. Die Restest-Reliabilitäten lagen bei $r_{12} = .66$, $r_{23} = .69$ und $r_{13} = .56$.

Depersonalisation (oder *Zynismus*) bezeichnet eine Verfassung, in der andere Personen nicht mehr als Mensch, sondern eher als Gegenstand, als Ding wahrgenommen und behandelt werden. Beispielsweise könnte ein Lehrer statt von seinen Schülern von seinem ‚Schülermaterial‘ sprechen. Ein Beispielim lautet: „Ich glaube, ich behandle Schüler zum Teil ziemlich unpersönlich“.

Die Reliabilitäten lagen zu den drei Meßzeitpunkten bei $\alpha = .69$, $.61$ und $.73$. Die Restest-Reliabilitäten lagen bei $r_{12} = .57$, $r_{23} = .62$ und $r_{13} = .55$.

Im Bereich der *Leistung*, die man im Beruf erbringt, hat man als ausgebrannter Mensch das Gefühl, nicht mehr soviel wie vorher bzw. nicht mehr genügend zu leisten, das Soll nicht mehr erfüllen zu können. Ein Beispielim lautet: „Ich glaube, daß ich das Leben anderer Menschen durch meine Arbeit positiv beeinflusse“ (–).

Die Reliabilitäten lagen zu den drei Meßzeitpunkten bei $\alpha = .82$, $.83$ und $.82$. Die Restest-Reliabilitäten lagen bei $r_{12} = .54$, $r_{23} = .74$ und $r_{13} = .50$.

Das Original-Antwortformat ist siebenstufig und reicht von „nie“, „mehrmals pro Jahr oder weniger“, „einmal im Monat“, „mehrmals im Monat“, „einmal die Woche“, „mehrmals die Woche“ bis „täglich“. Es wurde für die erste Datenerhebungswelle übernommen. Nach Protesten der Untersuchungsteilnehmer und einer relativ hohen Antwortverweigerung auf den drei Subskalen wurde ab der zweiten Welle das Antwortschema angeglichen und ebenfalls ein

vierstufiges Antwortformat statt des siebenstufigen gewählt (stimmt nicht (1) – stimmt kaum (2) – stimmt eher (3) – stimmt genau (4)).

Für die Daten des ersten Meßzeitpunktes wurde eine lineare Transformation vorgenommen, so daß alle Berechnungen in der vorliegenden Untersuchung auf dem vierstufigen Antwortformat beruhen.

In keiner der drei Wellen war es möglich, die dreifaktorielle Struktur des Burnout Inventars zu replizieren (vgl. Anhang B). Ebenso wenig konnte dann eine einheitliche Faktorstruktur über alle Erhebungswellen hinweg gefunden werden. Auch die Itemkennwerte sind zum Teil nicht befriedigend. Dennoch wurde auf eine Revision verzichtet, um die ursprünglichen Skalen beibehalten zu können, damit Vergleiche mit anderen Studien weiterhin möglich sind.

4.5 Stresseinschätzungen als Herausforderung, Bedrohung und Verlust

Die drei Subskalen zur Stresseinschätzung (Jerusalem, 1990) stellen Operationalisierungen im Sinne der transaktionalen Streß- und Emotionstheorie (Lazarus, 1991) dar. Danach können bedeutsame Ereignisse von Personen auf drei Arten kognitiv eingeschätzt werden: als Herausforderung, als Bedrohung oder (retrospektiv) als Schaden /Verlust. Die Skalen wurden in verschiedenen Studien empirisch geprüft (vgl. Jerusalem, 1990, 1993; Schwarzer, 1993a, 1993b).

Hoch selbstwirksame Personen neigen dazu, streßreiche Ereignisse als *Herausforderung* aufzufassen. Sie fühlen sich eher motiviert, angeregt und zuversichtlich im Hinblick auf die vor ihnen liegenden Aufgaben. Ein Item lautet: „Ich bin zuversichtlich, daß ich den Anforderungen im Beruf gewachsen bin“.

Die Reliabilitäten lagen zu den drei Meßzeitpunkten bei $\alpha = .62$, $.56$ und $.64$. Die Retest-Reliabilitäten lagen bei $r_{12} = .46$, $r_{23} = .61$ und $r_{13} = .54$.

Ein *Bedrohungserleben* bedeutet, daß eine Person die Befürchtung hegt, mit ihren verfügbaren Ressourcen und Kompetenzen die Anforderungen nicht bewältigen zu können, und somit auch entsprechend weniger zuversichtlich an die Aufgaben herangeht. Ein Item lautet: „Ich mache mir Sorgen wegen der Probleme, die auf mich zukommen“.

Die Reliabilitäten lagen zu den drei Meßzeitpunkten bei $\alpha = .79$, $.78$ und $.73$. Die Retest-Reliabilitäten lagen bei $r_{12} = .66$, $r_{23} = .61$ und $r_{13} = .55$.

Verlusterleben bezieht sich auf eine retrospektive Einschätzung von Ereignissen, die dadurch gekennzeichnet ist, daß die Person subjektiv einen Kontrollverlust empfindet; was sich in entsprechender Weise auch auf zukünftige Einschätzungen ähnlicher Anforderungssituati-

onen auswirken wird. Ein Beispielitem lautet: „Ich bin bedrückt, weil ich momentan einfach nicht klarkomme“.

Die Reliabilitäten lagen zu den drei Meßzeitpunkten bei $\alpha = .75$, $.74$ und $.70$. Die Restest-Reliabilitäten lagen bei $r_{12} = .47$, $r_{23} = .44$ und $r_{13} = .30$.

Das Antwortformat ist vierstufig: stimmt nicht (1) – stimmt kaum (2) – stimmt eher (3) – stimmt genau (4).

4.6 Berufliche Belastung

Die Skala wurde von Enzmann und Kleiber (1989) zur Erfassung der Belastung der konkreten beruflichen Tätigkeit entwickelt. Die Autoren stützten sich bei der Itemauswahl auf einen bereits früher entwickelten Fragebogen aus dem Modellversuch „Fachkräfte für die psychosoziale Versorgung“ der Freien Universität Berlin und der FHSS-Berlin (Kleiber & Enzmann, 1985).

Die Skala kann als Gesamtskala mit 15 Items eingesetzt werden. Die Reliabilitäten lagen zu den drei Meßzeitpunkten bei $\alpha = .83$, $.87$ und $.87$. Die Restest-Reliabilitäten lagen bei $r_{12} = .70$, $r_{23} = .75$ und $r_{13} = .58$.

Sie läßt sich aber auch in drei Subskalen untergliedern: (a) *Arbeitsüberforderung* (mit sechs Items wie „In meinem Beruf wird man ständig überfordert“). Die Reliabilitäten lagen zu den drei Meßzeitpunkten bei $\alpha = .73$, $.80$ und $.82$. Die Restest-Reliabilitäten lagen bei $r_{12} = .64$, $r_{23} = .71$ und $r_{13} = .56$.

(b) *Kontrolliertheitserleben* (mit drei Items wie „Ich habe das Gefühl, auch vor Kollegen ständig betonen zu müssen, daß ich viel arbeite“). Die Reliabilitäten lagen zu den drei Meßzeitpunkten bei $\alpha = .57$, $.63$ und $.85$. Die Restest-Reliabilitäten lagen bei $r_{12} = .54$, $r_{23} = .51$ und $r_{13} = .53$.

(c) *Arbeitsunzufriedenheit* (mit sechs Items wie „Meine eigenen Berufsideale lassen sich in meiner Tätigkeit nicht verwirklichen“). Die Reliabilitäten lagen zu den drei Meßzeitpunkten bei $\alpha = .81$, $.82$ und $.83$. Die Restest-Reliabilitäten lagen bei $r_{12} = .69$, $r_{23} = .72$ und $r_{13} = .59$.

Alle Items haben ein fünfstufiges Antwortformat (1 = trifft gar nicht zu, 2 = trifft wenig zu, 3 = trifft mittelmäßig zu, 4 = trifft überwiegend zu und 5 = trifft völlig zu).

4.7 Idealismus

Die zwei Items zum Idealismus im Lehrerberuf wurden einer Fünf-Item-Skala von Enzmann und Kleiber (1989) entnommen. Diese Skala von Enzmann und Kleiber stellt eine revidierte Fassung des „Konstanzer Fragebogens für Schul- und Erziehungseinstellungen“ (KSE) von Koch, Coetta und Müller-Fohrbrodt (1972) dar. Erfasst werden sollen vor allem die Ansprüche, die Lehrer in ihrer Berufsausübung an sich selber stellen (vgl. Enzmann & Kleiber, 1989).

Die Items lauten: „Wer in einer Tätigkeit wie meiner nicht mit Idealismus bei der Sache ist, sollte besser den Beruf wechseln“ und „Einen Beruf wie meinen sollte nur derjenige wählen, der bereit ist, um der Aufgabe willen auch beträchtliche persönliche Opfer zu bringen“.

Alle Items haben ein fünfstufiges Antwortformat (1 = trifft gar nicht zu, 2 = trifft wenig zu, 3 = trifft mittelmäßig zu, 4 = trifft überwiegend zu und 5 = trifft völlig zu).

Die Reliabilitäten lagen zu den drei Meßzeitpunkten bei $\alpha = .55$, $.59$ und $.65$. Die Restest-Reliabilitäten lagen bei $r_{12} = .62$, $r_{23} = .63$ und $r_{13} = .54$.