

Empirische Sonderpädagogik, 2020, Nr. 1, S. 27-44
ISSN 1869-4845 (Print) · ISSN 1869-4934 (Internet)

Entwicklung der Kurzskala zur Selbstwirksamkeit im Umgang mit heterogenen Lerngruppen (KUS-HL) – Reanalysen der Skalen von Kopp (2009) und Validierung der Kurzskala

Steffen Siegemund¹, Michel Knigge² & Carolin Rotter³

¹Universität Würzburg, ²Universität Potsdam, ³Universität Duisburg-Essen

Zusammenfassung

Die vorliegende Studie überprüft die Faktorenstruktur der Selbstwirksamkeitsskalen im Umgang mit Heterogenität bzw. mit „besonderen“ Schülerinnen und Schülern (SuS) von Kopp (2009) anhand von zwei Datensätzen mit insgesamt 725 Studierenden der Lehrämter Primar- und/oder Sekundarstufe, Gymnasium sowie Sonderpädagogik. Im Ergebnis zeigt sich ein besserer Modellfit für ein alternatives vierfaktorielles Modell. Die Aufteilung in die Faktoren „Adaptive Unterrichtsgestaltung in heterogenen Lerngruppen“ und „Adaptive Unterrichtsgestaltung für SuS mit erhöhtem sonderpädagogischen Förderbedarf“ verweist hierbei auf die viel diskutierte Unterscheidung eines engen und eines weiten Inklusionsbegriffs (vgl. Werning, 2014) auch auf Ebene der Selbstwirksamkeit von Studierenden. Die weiteren Faktoren „Inklusive Lehrerpersönlichkeit“ und „Klassenklima“ folgen der Struktur von Kopp, wenngleich die Neuordnung einiger Items eine stringenter Interpretation ermöglicht und den Modellfit verbessert. An einem dritten Datensatz ($n = 456$) erfolgt die Validierung der neu entwickelten Kurzskala KUS-HL. Diese bietet zusammenfassend eine ökonomische Alternative zur Messung der Selbstwirksamkeit auch für Gruppen von Studierenden mit unterschiedlichen Lehrämtern.

Schlüsselwörter: Inklusion, Heterogenität, Überzeugung, Selbstwirksamkeitserwartung, Lehramtsstudierende

Development of the short scale for self-efficacy in dealing with heterogeneous learning groups (KUS-HL) – Reanalyses of the scales by Kopp (2009) and validation of the short scale

Abstract

This study examines the factor structure of self-efficacy scales dealing with heterogeneity or „special“ pupils developed by Kopp (2009). This is conducted through two sets of data with a total number of 725 students pursuing a teaching degree in the fields of primary and/or secondary school or special needs education. The results show a better model fit for an alternative four-factor model. The division into the factors „Adaptive teaching structure in heterogeneous learning groups“ and „Adaptive teaching structure for pupils with increased special educational needs“ refers to the much discussed distinction between a narrow and a broad concept of inclusion (cf. Werning, 2014), also at the level of students' self-efficacy. The other factors, „Inclusive

teacher personality“ and „Class atmosphere“, follow Kopp’s structure, although the reorganization of some items allows a more stringent interpretation and improves the model fit. A third set of data ($n = 456$) is used to validate the newly developed short scale KUS-HL. In summary, this scale offers an economic alternative to measure self-efficacy even for groups of students with differing teaching degrees.

Keywords: Inclusive education, heterogeneity, beliefs, self efficacy

Theoretischer Hintergrund

Selbstwirksamkeit als Disposition

Die Ausgestaltung von Rahmenmodellen zur Beschreibung der professionellen Kompetenz von Lehrkräften im Hinblick auf die spezifischen Anforderungen schulischer Inklusion stellt aktuell eine wichtige Herausforderung innerhalb der Erziehungswissenschaft dar (vgl. Heinrich, Urban & Werning, 2013; Kopp, 2009; Melzer, Hillenbrand, Sprenger & Hennemann, 2015). Das Modell „Kompetenz als Kontinuum“ integriert personenbezogene Dispositionen, situationsspezifische Fähigkeiten zur Handlungsvorbereitung und die beobachtbare Performanz (vgl. Blömeke, Gustafsson & Shavelson, 2015). Im Bereich der Dispositionen interagieren Professionswissen, Überzeugungen und affektiv-motivationale Orientierungen. Einstellungen sowie Selbstwirksamkeitserwartungen (im Weiteren als „Selbstwirksamkeit“ bezeichnet) zu inklusiven schulischen Lehr- und Lernsettings bilden aktuell bei der genannten Ausgestaltung einen zentralen Forschungsbereich. Selbstwirksamkeit drückt die Zuversicht aus, eine herausfordernde Situation bewältigen zu können. Gerade im Bereich schulischer Inklusion werden diese in vielfältiger Art und Weise wahrgenommen. Entsprechend wird Selbstwirksamkeit zur Bewältigung dieser Herausforderungen als eine Gelingensbedingung schulischer Inklusion vielfach benannt (vgl. Bosse & Spörer, 2014; Kopmann & Zeins, 2016).

Facetten von Selbstwirksamkeit

Im Sinne der sozial-kognitiven Theorie von Bandura (1997) bezeichnet Selbstwirksamkeit die subjektive Erwartung, eine schwierige Anforderungssituation bewältigen zu können. Zwei Items zur Messung einer spezifischen Selbstwirksamkeit von Lehrkräften finden sich erstmalig in Evaluationsprojekten der RAND Corporation in den 1970er Jahren. Das erste Item, „When it comes right down to it, a teacher really can’t do much because most of a student’s motivation and performance depends on his or her home environment“, fokussiert dabei auf grundsätzliche Annahmen zur Beeinflussbarkeit von psychischen oder sozialen Systemen (zitiert nach Schmitz & Schwarzer, 2002, S. 193). Das zweite Item entspricht hingegen stärker dem Konzept von Bandura, da ein positives Ergebnis durch Anstrengung und Abrufen der eigenen Kompetenzen formuliert wird: „If I try really hard, I can get through to even the most difficult or unmotivated students“. Die eher „unpräzise“ Formulierung dieses Items unterstützt dabei das Erfragen einer allgemeinen Zuversichtlichkeit angesichts schwieriger pädagogischer Situationen. Das verwandte Konstrukt der Handlungsergebnis-Erwartung beschreibt hingegen einen angenommenen Kausalzusammenhang zwischen sachgerichteten Handlungen und den erzielten Handlungsergebnissen (vgl. Schmitz & Schwarzer, 2000).

In ihrer Selbstwirksamkeitsskala unterscheiden Schmitz und Schwarzer (2000) vier relevante Bereiche der beruflichen Tätigkeit von LehrerInnen: (a) allgemeine be-

rufliche Leistungen, (b) berufsbezogene soziale Interaktion, (c) innovatives pädagogisches Handeln und (d) Umgang mit Stress und Emotionen. Die Erkenntnis der Bereichsspezifität von Selbstwirksamkeit führte zu weiteren fachspezifischen Differenzierungen des Konstrukts in weitere Faktoren bzw. Facetten (vgl. Gebauer, 2013).

Einflussfaktoren auf Selbstwirksamkeit

Enge Zusammenhänge von Einstellungen und Selbstwirksamkeit in Bezug auf schulische Inklusion werden mehrfach berichtet (vgl. Bosse et al., 2016; Hellmich & Görel, 2014; Kopmann & Zeins, 2016). Für die Selbstwirksamkeit im inklusiven Unterricht konnten Zusammenhänge zu den eigenen Erfahrungen mit SuS mit sonderpädagogischen Förderbedarfen als auch Erfahrungen mit inklusiven Lehr- und Lernsettings nachgewiesen werden (Dessemontet, Morin & Crocker, 2014; Hellmich, Görel & Schwab, 2016; Malinen et al., 2013). Wenngleich vielfach die hohe Stabilität von Einstellungen und Selbstwirksamkeit herausgestellt wird (vgl. Woodcok, Hemmings & Kay, 2012), liegen dennoch Hinweise auf die Wirksamkeit solcher Lerngelegenheiten vor (z.B. De Boer, Pijl & Minaert, 2011; Weisel & Dror, 2006). Zusammenhänge von Lehrerfortbildungen und der Selbstwirksamkeit von Lehrkräften fanden Richter, Engelbert, Weirich und Pant (2013) sowie Wudy und Jerusalem (2011). Erste Befunde liegen vor, nach denen auch inklusionsspezifische Selbstwirksamkeit durch Trainings gestärkt werden kann (Sharma & Nuttal, 2015; Forlin, Loreman & Sharma, 2014).

Effekte der Selbstwirksamkeit

Für eine hohe Selbstwirksamkeit von Lehrkräften in der Emotionsregulation können

als positive Wirkungen ein niedrigeres Stressempfinden und eine höhere Arbeitszufriedenheit benannt werden (vgl. Abele & Candova, 2007; Krause, Dorsemagen & Baeriswyl, 2013; Wudy & Jerusalem, 2011). Positive Outcomes auf SuS werden vielfach in der Literatur benannt. Knigge und Rotter (2015) konnten z.B. zeigen, dass Studierende mit einer höheren Selbstwirksamkeit in ihren Überlegungen zu einer Unterrichtsplanung größeren Wert auf die Motivierung aller SuS legen. Unterrichtsbezogene Selbstwirksamkeit bedingt auch den Einsatz lernförderlicher Unterrichtsmaßnahmen, wie z.B. die kognitive Aktivierung (Depaepe & König, 2018). Der Zusammenhang von Selbstwirksamkeit und Performanz als Lehrkraft in inklusiven Settings nach Maßgabe des Lernerfolgs der SuS ist aktuell noch unzureichend erforscht.

Ausgangspunkt

Im Rahmen des Hamburger Projekts *ProfaLe*¹ werden u.a. verschiedene Instrumente zur Messung von Kompetenzen mit Bezug zur schulischen Inklusion erprobt und weiterentwickelt. Zur Evaluation von neu entwickelten Seminarkonzepten und zur Erprobung der Instrumente wurden im Sommersemester 2016 neben Skalen zur Einstellung zur schulischen Inklusion auch die von Kopp (2009) entwickelten Selbstwirksamkeitsskalen eingesetzt. Die Auswahl der Skalen orientierte sich an der *EiLink-Studie* (Knigge & Rotter, 2014).

Da im Handlungsfeld Inklusion (*ProfaLe*) die Verknüpfung sonderpädagogischer und fachdidaktischer Perspektiven auf Unterrichtsprozesse in inklusiven Klassen ein zentrales Anliegen darstellt und die Skala von Kopp Herausforderungen im inklusiven Unterricht fokussiert, ist inhaltlich eine

¹ Das Projekt *Professionelles Lehrerhandeln zur Förderung fachlichen Lernens unter sich verändernden gesellschaftlichen Bedingungen (ProfaLe)* der Universität Hamburg wird im Rahmen der gemeinsamen „Qualitätsoffensive Lehrerbildung“ von Bund und Ländern aus Mitteln des Bundesministeriums für Bildung und Forschung gefördert.

hohe Passung gegeben. Allerdings werfen sowohl das methodische Vorgehen bei der Skalenkonstruktion als auch die inhaltliche Interpretation der vier identifizierten inklusionsspezifischen Faktoren Fragen auf.

Bei methodischen Vorgehen kann die Verwendung des Varimax Rotationsverfahrens der EFA, welches orthogonale, d.h. unkorrelierte Faktoren erzeugt, kritisch beurteilt werden. Zunächst werden empirische Daten zumeist besser durch korrelierte Faktoren abgebildet (vgl. Costello & Osborne, 2005). Dieses gilt insbesondere für hochspezifische Konstrukte, zu denen die Selbstwirksamkeit als Lehrkraft in inklusiven Schulen mit dem Fokus auf der Interaktion im Klassenraum (unter Ausschluss von Schulklima, Elternarbeit etc.) gezählt werden kann. Bei der Interpretation der gebildeten Faktoren konnte die Zugehörigkeit zum jeweiligen Faktor für mehrere Items nicht nachvollzogen werden, sodass sich auch Fragen im Hinblick auf die inhaltliche Definition der einzelnen Faktoren stellen.

Ein weiteres Argument für die Überprüfung der von Kopp gefundenen Faktorenstruktur leitet sich aus der Beschränkung der Stichprobe auf Studierende des Lehramts für Grundschulen ab. Ob das gefundene Modell auch zu Stichproben von anderen Lehramtsstudierenden passt, ist eine empirische Frage, die im vorliegenden Artikel beantwortet werden soll.

Fragestellung

Dieser Beitrag prüft zunächst (Fragestellung 1) durch konfirmatorische Faktorenanalysen (CFA), ob das Modell von Kopp zu einer Stichprobe von Studierenden der Lehramter Sonderpädagogik, Primar- und Sekundarstufe sowie Gymnasium passt.

Fragestellung 2 sondiert mit Hilfe explorativer Analysen alternative Faktorenstrukturen mit einer jeweils möglichst stringenten Interpretation der Faktoren und Passung der Items. Ergänzend erfolgt eine statistisch ge-

lenkte Reduktion der Items zur Konstruktion einer Kurzskaala mit 12 Items.

Erneute CFA sollen die Frage beantworten, ob die neue Faktorenstruktur (als vollständiges Set mit 24 sowie in der Kurzform mit 12 Items) die Selbstwirksamkeit von Studierenden verschiedener Lehramter in Bezug auf „besondere“ SuS mit einem guten Fit modellieren kann (Fragestellung 3). Da Inklusion als Querschnittsaufgabe zu verstehen ist und somit Studierende und Lehrkräfte aller Lehramtsformen involviert sind, stellt sich ergänzend die Frage, ob die gebildeten Skalen für die verschiedenen Lehramter das gleiche Konstrukt in gleicher Ausprägung erfassen (Prüfung der Messinvarianz).

Methode

Stichprobe

Die Gesamtstichprobe für alle 24 Items von Kopp beinhaltet 725 Datensätze von Studierenden der Universitäten Hamburg und Halle, die mindestens an einem Messzeitpunkt teilgenommen haben (siehe Tab.1). Die (Teil-)Stichprobe 1 wird im Folgenden für alle CFA der Langform mit 24 Items verwendet. Sie besteht aus 565 Studierenden, die an einer zu Beginn eines Semesters liegenden Erhebungswelle teilgenommen haben. Die (Teil-)Stichprobe 2 kommt in den explorativen Faktorenanalysen (EFA) zur Anwendung und besteht aus 384 Studierenden, die an einer am Ende eines Semesters liegenden Erhebungswelle teilgenommen haben. Für 224 von diesen Studierenden liegen Längsschnittdaten vor. Die verwendeten Daten aus der EiLink-Studie wurden im WiSe 2013/14 und die Daten aus dem Projekt Profale im SoSe 2016 mit Hilfe eines Onlinefragebogens erhoben.

Die neue Kurzskaala KUS-HL wurde zusätzlich an 456 Studierenden der Europa-Universität Flensburg und den Universitäten Erfurt und Hamburg an einem Erhebungszeitpunkt erprobt (Stichprobe 3).

Tabelle 1: Deskriptive Statistiken zu den Stichproben 1, 2 und 3

Zeitpunkt/ Stichprobe	Ge- samt	Studienabschnitt				Lehramt			
		Bachelor		Master	n.a.	Primar und/oder Sek 1	Gym./ Sek 2	Son- der- päd.	n.a.
		1-3	4-6						
1 (Anfang Semester)	565 100%	94 16,6%	211 37,3%	250 44,2%	10 1,8%	278 49,2%	186 32,9%	99 17,5%	2 0,4%
2 (Ende Semester)	384 100%	82 21,4%	127 33,1%	169 44,0%	6 1,6%	184 47,9%	112 29,2%	79 20,6%	9 2,3%
3 (Kurz)	456 100%	156 34,1%	225 49,2%	64 14%	12 2,6%	238 52,1%	18 3,9%	168 36,8%	33 7,2%

Anmerkungen: Im Lehramt „Primar und/oder Sek 1“ sind verschiedene Lehrämter der beteiligten Universitäten zusammengefasst, z.B. das Lehramt „Grundschule“.

Variablen

Der Fragebogen von Kopp (2009) umfasst die folgenden vier Faktoren: „Adaptive Unterrichtsgestaltung“ (1), „Stiftung eines inklusiven Klassenklimas“ (2), „Inklusive Lehrerpersönlichkeit“ (3) und „Erfolgreiches Unterrichten“ (4). Die Konstruktionsprinzipien der 24 Items und theoretische Annahmen zu möglichen Faktoren sind nicht publiziert. Kopp achtet bei der Formulierung der Items besonders darauf, verschiedene Heterogenitätsdimensionen in den Blick zu nehmen. Neben der Benennung verschiedener sonderpädagogischer Förderbedarfe finden sich Formulierungen, die „Besonderheiten“ in den Feldern „Migrationshintergrund“, „Verhaltensauffälligkeiten“ und „Leistungsheterogenität“ ausdrücken. Die damit verbundene Annahme, nach der die verschiedenen Differenzlinien sich auch in der Faktorenstruktur niederschlagen würden, konnte allerdings nicht durch die empirischen Befunde bestätigt werden. In der Selbstwirksamkeit im Umgang mit „Heterogenität im Lernen und Leisten“ zeigte sich – entgegen der Erwartung von Kopp – kein Einfluss der expliziten Adressierung von SuS mit Behinderung (vgl. Kopp, 2009).

Die Bildung der Faktoren erfolgte induktiv auf der Grundlage einer explorativen Hauptkomponentenanalyse mit orthogonaler Varimax-Rotation. Eine konfirmatorische

Überprüfung der Faktorenstruktur an sich sowie unter Einbezug einer Stichprobe verschiedener allgemeinbildender Lehrämter liegt bislang nicht vor.

Der Fragebogen wurde bereits mehrfach in weiteren Studien entweder im Ganzen verwendet (Knigge & Rotter 2015) oder es wurden alternativ einzelne Items bzw. Faktoren übernommen (Bosse & Spörer, 2014; Hellmich & Görel, 2014; Koppmann & Zeins, 2016).

Die Zuordnung einzelner Items zu den vier genannten Faktoren wirft – wie bereits berichtet – Fragen auf, die zur Überprüfung der Faktorenstruktur motiviert haben. So fällt z.B. bei der Zuordnung auf, dass Faktor 4 „Erfolgreiches Unterrichten“ nur Items beinhaltet, die negativ formuliert sind, d.h. rekodiert werden müssen, und zusätzlich inhaltlich recht heterogen sind (vgl. Tab. 3). So beinhaltet Item 5 zwar das Wort „Leistung“ als Indikator für einen Lernerfolg, aber die Selbstwirksamkeit ist nicht auf die Leistungsentwicklung gerichtet, sondern auf die eigene Haltung gegenüber der Leistung der SuS: *„Ich glaube, dass ich es auch beim besten Willen nicht schaffen werde, die Leistungen aller Kinder gleichermaßen zu würdigen und ihren Voraussetzungen entsprechend zu beurteilen, selbst wenn sie das Lernziel nicht erreicht haben“*. Ebenso zielt Item 11 auf die Wertschätzung der Leistung,

allerdings diesmal nicht durch die Lehrkraft, sondern durch die MitschülerInnen.

Weitere Items, die die Würdigung von Leistungen durch die Lehrkraft berücksichtigen, finden sich in Faktor 3 „Inklusive Lehrerpersönlichkeit“. Während das hierzu passende Item 5 ausgeschlossen wurde, findet sich dafür in Faktor 3 das eher zu Faktor 1 „Adaptive Unterrichtsgestaltung“ passende Item 9: *„Ich weiß, dass ich auch in Klassen mit großen Leistungsunterschieden die Lernvoraussetzungen eines jeden Kindes erkennen kann und meinen Unterricht darauf aufbauend planen kann“*. In Faktor 2 „Inklusives Klassenklima“ finden sich mehrere Items, die sich auf die Anerkennungsverhältnisse zwischen den SuS beziehen, allerdings auch das Item 3: *„Ich fühle mich in der Lage, den Bedürfnissen aller Kinder gerecht zu werden und beispielsweise ein Kind mit starker körperlicher Beeinträchtigung zu pflegen.“*

Zusammenfassend erscheint die Unterscheidung des Bereichs „Anerkennungsverhältnisse und Wertschätzung“ in die Beziehungen Schüler-Schüler (Faktor 3) und Lehrkraft-Schüler (Faktor 4) als trennscharf und plausibel. Die Zuordnung ist allerdings nicht für alle Items nachvollziehbar. Die Abgrenzung dieses Bereichs zu einem weiteren, der die Leistungsentwicklung beinhaltet, lässt sich ebenfalls mit theoretischen Überlegungen gut in Einklang bringen (vgl. Grosche, 2015). Dieser zweite Bereich fokussiert die fachbezogene Leistungsentwicklung der SuS und passt zu Items aus Faktor 1 „Adaptive Unterrichtsgestaltung“ und Faktor 4 „Erfolgreiches Unterrichten“. Bei der Bezeichnung verweist Faktor 1 stärker auf das Handeln der Lehrkraft und Faktor 4 stärker auf die Auswirkungen auf Seite der SuS. Da das Konzept der Selbstwirksamkeit aber eben diese Beziehung, d.h. die Wirkung der Lehrkraft auf die SuS erfassen soll, verbleibt konzeptuell und auf Ebene der Itemformulierungen diese Unterscheidung unklar.

Konfirmatorische Überprüfung Modell „Kopp“

Analysemethode CFA

Die Überprüfung der Faktorenstruktur von Kopp erfolgt anhand des Datensatzes 1 ($n = 565$). Die Analyse erfolgt mit der Software Mplus (8.1.5). Für den Umgang mit fehlenden Werten kommt hierbei das „Full Information Maximum Likelihood“-Verfahren (FIML) zum Einsatz. Für die Beurteilung der MCAR-Bedingung wurde der Test nach Little angefordert. Im Ergebnis muss die Nullhypothese verworfen werden, d.h. MCAR-abhängige Verfahren zur Behandlung fehlender Werte (z.B. listenweiser Ausschluss) sind unzulässig (vgl. Baltes-Götz, 2013).

Für alle Items wurde die Stichprobe 1 zusätzlich auf Normalverteilung getestet. Der Kolmogorov-Smirnov-Test auf Normalverteilung ist für alle Items signifikant ($p < .01$), d.h. die Nullhypothese wird abgelehnt und die Daten sind nicht normalverteilt. Als Konsequenz wurde im Weiteren der robustere MLR-Schätzer verwendet (vgl. Kline, 2015).

Ergebnisse CFA Modell „Kopp“

Der Fit für das unkorrelierte Modell ist in allen Indizes nicht akzeptabel (vgl. Tab. 2). Für den „Overall Modellfit“, berechnet mit dem SRMR, gilt als akzeptabler Fit z.B. $SRMR \leq .10$ (vgl. Schermelleh-Engel, Moosbrugger & Müller, 2003; Little, 2013). Diese Marke verfehlt das unkorrelierte Modell mit einem SRMR von .230 deutlich. Das Modell „Kopp oblique“ mit korrelierten Faktoren zeigt hingegen gute bis akzeptable Werte in den Overall-Fit-Indizes RMSEA (.049-.059; guter Fit: $RMSEA < .05$) und SRMR (.053; akzeptabler Fit $.05 < SRMR \leq .10$). Bei den inkrementellen Indizes CFI (.881) und TLI (.867) wird allerdings kein akzeptabler Fit erzielt (akzeptabler Fit: $.97 > CFI/TLI \geq .95$; vgl. Schermelleh-Engel et al., 2003).

Tabelle 2: Anpassungsgüte des Kopp-Modells. Stichprobe 1 ($n = 565$)

Faktorenmodell	Anzahl Items	χ^2	df	χ^2/df	CFI	TLI	RMSEA (90%CI)	SRMR
Kopp orthogonal	24	1392.81	252	5.53	.675	.644	.084-.093	.230
Kopp oblique	24	662.37	246	2.70	.881	.867	.049-.059	.053

Anmerkungen: Im Lehramt „Primar und/oder Sek 1“ sind verschiedene Lehrämter der beteiligten Universitäten zusammengefasst, z.B. das Lehramt „Grundschule“.

Fragestellung 1, die danach fragt, ob das Modell von Kopp zu einer Stichprobe von Studierenden der Lehrämter Sonderpädagogik, Primar- und Sekundarstufe sowie Gymnasium passt, muss auf Grund eines schlechten Modellfits für das orthogonale Modell verneint werden. Die Alternative mit korrelierten Faktoren bietet hingegen einen akzeptablen bis mittelmäßigen Fit.

Explorative Analysen zur Entwicklung einer alternativen Faktorenstruktur

Analysemethode EFA

Anhand der Stichprobe 2 ($n = 384$) soll durch EFA untersucht werden, ob eine alternative Faktorenstruktur gefunden werden kann, die zunächst eine kohärentere Interpretation der gebildeten Faktoren ermöglicht und im nächsten Schritt auch einen besseren Modellfit aufweist (Fragestellung 2). Die Analyse wurde mit Mplus durchgeführt. Für die EFA kommt ein Maximum-Likelihood-Schätzer zum Einsatz. Unter der Annahme korrelierter Faktoren wird Promax als Rotationsmethode verwendet.

Mit Hilfe des Kaiser-Guttman-Kriteriums, eines Scree-Tests und einer Parallelanalyse wird erneut die Frage nach der Anzahl der zu extrahierenden Faktoren aufgegriffen (vgl. Moosbrugger & Schermelleh-Engel, 2012). Die Parallelanalyse wird berechnet mit dem Programm „Monte Carlo PCA for Parallel Analysis Version 2.3“ (Watkins, 2008; vgl. Watkins, 2006).

Ergebnisse EFA

Mit einem KMO-Wert von .91 kann den Items zunächst eine hohe Zusammengehörigkeit attestiert werden (vgl. Kaiser & Rice, 1974). Dies bestätigt sich auch über den mit $p < .001$ signifikanten Bartlett-Test auf Sphärizität, sodass wichtige Voraussetzungen für eine EFA gegeben sind.

Abbildung 1 stellt die Ergebnisse der EFA im Hinblick auf die Frage nach der Anzahl der zu bildenden Faktoren dar. Die Linie mit den Dreiecken als Markierung ist der Screeplot der empirischen Daten. Ein Knick zeigt sich beim fünften Faktor, welcher auf vier Faktoren im Einklang mit dem Modell von Kopp hindeutet. Ein leichter Knick zeigt sich zusätzlich beim dritten Faktor sowie ein deutlicher beim zweiten. Der Paralleltest zeigt ebenfalls zwei Faktoren mit einem Eigenwert über den Zufallsfaktoren. Die Faktoren 3 und 4 erreichen zumindest das Niveau der Zufallsdaten und liegen außerdem deutlich über der Grenze des Kaiser-Guttman-Kriteriums, das fünf Eigenwerte größer als 1 (7.92, 1.66, 1.39, 1.33, 1.02) herausstellt (vgl. Wentura & Pospechill, 2015).

Zusammenfassend legen die EFA ein zwei- oder vierfaktorielles Modell nahe. Der Vergleich der beiden Ladungsmatrizen zeigt eine genestete Relation der beiden Modelle, sodass diese im Weiteren direkt gegeneinander getestet werden können. Die genestete Relation kann auch inhaltlich interpretiert werden. Im zweifaktoriellen Modell bilden Items den ersten Faktor, die sich auf die Selbstwirksamkeit im Umgang mit einzelnen SuS beziehen. Drei Items, die

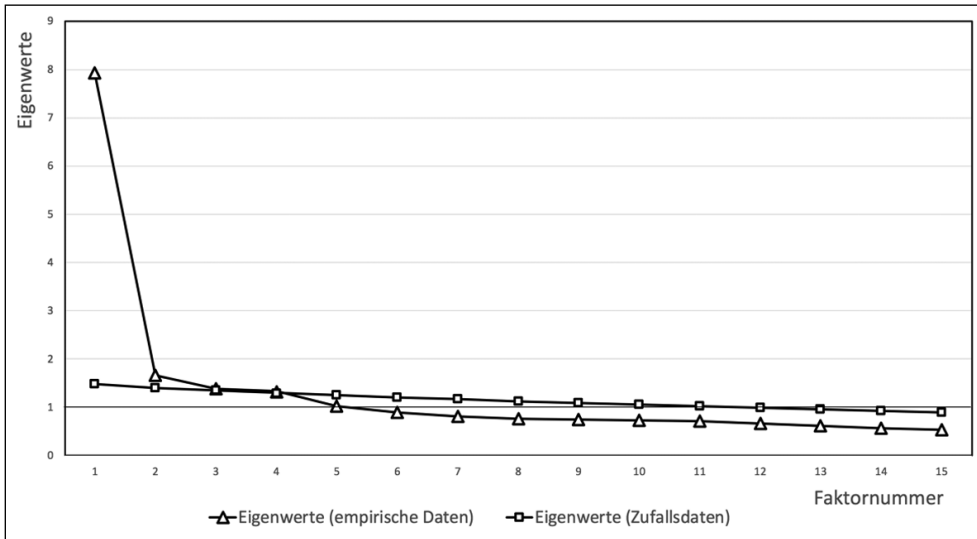


Abbildung 1: Screeplot und Parallelanalyse. Stichprobe 2 ($n = 384$)

sich auf das Klassenklima beziehen, bilden den zweiten Faktor.

Das vierfaktorielle Modell unterscheidet drei Faktoren (1–3) mit Bezug auf einzelne SuS und einen zum Klassenklima (4):

1. Der erste Faktor vereint alle Items, die sich auf die Passung des Unterrichts für einzelne SuS und deren Leistungsentwicklung beziehen (siehe Tabelle 3). Fünf der insgesamt acht Items finden sich auch im Faktor 1 „Adaptive Unterrichtsgestaltung in heterogenen Lerngruppen“ vom Ursprungsmodell (Kopp, 2009). Auch Item 9 (zuvor in Faktor 3) „Ich weiß, dass ich auch in Klassen mit großen Leistungsunterschieden die Lernvoraussetzungen eines jeden Kindes erkennen und meinen Unterricht darauf aufbauend planen kann“ findet sich nun in Faktor 1. Die höchste Faktorladung weist das Item 25 „Ich bin mir sicher, so vielfältiges Lernmaterial anbieten zu können, dass auch Kinder mit Sprachschwierigkeiten im Sachunterricht folgen können“ auf. Mit dem Begriff „Sprachschwierigkeiten“ sind vermutlich SuS mit Deutsch als Zweit- oder Fremdsprache adressiert. Sonderpädagogische Förderbedarfe werden nicht

explizit benannt. Inhaltlich passen auch die Items 21 und 12, die aber jeweils nur die zweithöchste Ladung auf Faktor 1 haben. Diese Items sind allerdings durch ihre Länge und ihre Umpolung komplex. Inhaltlich ist zudem jeweils unklar, ob der Lernerfolg oder die Beziehung zu den MitschülerInnen (ausgedrückt durch die Wörter „mitlernen“ und „nicht stören“) im Vordergrund steht.

2. Fünf von insgesamt sechs Items aus Faktor 2 gehören im Ursprungsmodell ebenfalls zum Faktor 1, sodass dieser aufgespalten wird. Der in den Reanalysen extrahierte Faktor 2 beinhaltet alle Items, die sich explizit auf das Unterrichten von SuS mit (erhöhtem) sonderpädagogischem Förderbedarf bzw. mit Behinderungen beziehen. So benennen Item 3 und 7 körperlich(schwerst)-behinderte Kinder als Zielgruppe, Item 16 „gehörlose oder taubstumme Kinder“ sowie Item 2, 14 und 20 geistig behinderte Kinder bzw. solche mit geistiger Entwicklungsverzögerung. Die höchste Ladung auf Faktor 2 „Adaptive Unterrichtsgestaltung für SuS mit erhöhtem sonderpädagogischen Förderbedarf“ hat

Tabelle 3 (diese und folgende Seite): Ladungsmatrix und Trennschärfe für alle 24 Items der neu entwickelten Faktorenstruktur der Skala "Selbstwirksamkeit im Umgang mit besonderen Kindern" im Ergebnis EFA mit dem obliquen Rotationsverfahren Promax anhand des Datensatzes 2 (n = 384)

Faktor	1	2	3	4	Faktor	r
Itemzahl	8	6	5	5	Modell	
Cronbachs alpha	.775	.828	.771	.621	Kopp	
Skalenmittelwert	2.79	2.48	3.05	2.94		
Faktor 1: Adaptive Unterrichtsgestaltung in heterogenen Lerngruppen						
13 „Ich fühle mich in der Lage, Unterrichtsarrangements so offen zu gestalten, dass ich mich um Kinder mit besonderen sprachlichen Schwierigkeiten gezielt kümmern kann, auch wenn andere Kinder bereits viel weiter fortgeschritten sind.“	.650	.455	.384	-	1	.574
22 „Ich kann Unterricht auch im bestehenden System so organisieren, dass sogar hochbegabte Kinder in ihrem Stärken gefördert werden können.“	.443	-	-	-	1	.302
9 „Ich weiß, dass ich auch in Klassen mit großen Leistungsunterschieden die Lernvoraussetzungen eines jeden Kindes erkennen und meinen Unterricht darauf aufbauend planen kann.“	.573	.401	.332	-	3	.563
17 „Ich bin mir sicher, dass ich auch bei größten Leistungsunterschieden für jedes Kind ein angemessenes Lernangebot bereithalten kann.“	.671	.514	.436	-	1	.617
1 „Ich bin mir sicher, dass ich auch in Regelklassen Unterricht so organisieren kann, dass leistungsstärkere auch von leistungsschwächeren Kindern profitieren können.“	.572	.307	.347	.313	1	.473
25 „Ich bin mir sicher, so vielfältiges Lernmaterial anbieten zu können, dass auch Kinder mit Sprachschwierigkeiten im Sachunterricht folgen können.“	.682	.458	.537	.525	1	.554
21 „Wenn ich mich auch noch so sehr bemühe, ich werde in meiner Klasse den Unterricht in Deutsch nicht so gestalten können, dass auch Kinder mit unzureichenden Sprachkenntnissen erfolgreich mitlernen können.“ *	.301	-	(.456)	-	3	.314
12 „Auch wenn ich Unterricht offen und individualisierend gestalte, wird es mir nicht gelingen, Kinder mit problematischem Verhalten über eine längere Zeit hinweg zu konzentrierter Arbeit, die die anderen nicht stört, zu motivieren.“ *	.342	.285	(.466)	.324	4	.410
Faktor 2: Adaptive Unterrichtsgestaltung für SuS mit erhöhtem sonderpädagogischem Förderbedarf						
7 „Ich kann Unterricht methodisch so gestalten, dass auch körperlich schwerstbehinderte Kinder ein für sie angemessenes Arbeitsangebot und einen passenden Arbeitsrhythmus finden können.“	.452	.686	.359	.373	1	.622
16 „Es gelingt mir, Unterricht so zu organisieren, dass auch gehörlose oder taubstumme Kinder genügend Gelegenheit finden, mit mir und den Mitschülern/innen in Kontakt zu treten.“	.359	.567	-	-	1	.562
2 „Selbst wenn ich mich sehr bemühe, wird es mir nicht gelingen, auf die Lernvoraussetzungen eines geistig behinderten Kindes Rücksicht zu nehmen und dementsprechend meinen Unterricht zu organisieren.“ *	-	.577	.465	.307	1	.589

Faktor	1	2	3	4	Faktor	r
Itemzahl	8	6	5	5	Modell	
Cronbachs alpha	.775	.828	.771	.621	Kopp	
Skalenmittelwert	2.79	2.48	3.05	2.94		
20 „Ich traue mir zu, Unterricht so zu organisieren, dass auch Kinder mit geistiger Entwicklungsverzögerung in ihrem eigenen Lerntempo zum Ziel kommen können.“	.600	.615	.552	.466	1	.633
14 „Ich weiß, dass ich ein Unterrichtsthema so vielfältig aufbereiten kann, dass auch Kinder mit geistigem Förderbedarf aktiv am Unterricht teilnehmen können.“	.550	.672	.392	.345	1	.625
3 „Ich fühle mich in der Lage, den Bedürfnissen aller Kinder gerecht zu werden und beispielsweise ein Kind mit starker körperlicher Beeinträchtigung zu pflegen.“	.388	.750	.426	.312	2	.572
Faktor 3: Inklusive Lehrerpersönlichkeit – Anerkennung aller SuS						
8 „Es wird mir schwerfallen, auch die Arbeiten jener Kinder, die ständig meinen Unterricht stören, zu würdigen und gerecht zu bewerten.“ *	-	-	.552	-	3	.370
4 „Ich bin mir sicher, dass ich auch jene Schüler/innen, die mir im Unterricht durch Störungen Probleme bereiten, als Person annehmen und schätzen kann.“	.436	.368	.701	.424	3	.445
10 „Selbst wenn der Unterricht durch verhaltensauffällige Schüler/innen gestört wird, bin ich mir sicher, die notwendige Gelassenheit bewahren zu können.“	.428	.318	.529	-	3	.386
19 „Ich bin in der Lage, die Leistungen aller Kinder ihren Lernvoraussetzungen gemäß zu würdigen, so dass auch ein Kind mit fremder Muttersprache stets einen Lernerfolg erfährt.“	.569	.352	.582	.502	3	.314
5 „Ich glaube, dass ich es auch beim besten Willen nicht schaffen werde, die Leistungen aller Kinder gleichermaßen zu würdigen und ihren Voraussetzungen entsprechend zu beurteilen, selbst wenn sie das Lernziel nicht erreicht haben.“ *	.325	-	.550	-	4	.361
Faktor 4: Inklusives Klassenklima						
24 „Ich kann es schaffen, die Klassengemeinschaft so zu beeinflussen, dass sich auch Kinder mit schwersten körperlichen Behinderungen darin angenommen fühlen.“	.384	.391	.459	.858	2	.653
23 „Ich werde es schaffen, dass auch Kinder mit geistiger Behinderung von den Mitschülern/innen akzeptiert werden.“	.425	.321	.464	.801	2	.627
18 „Ich kann das Klassenklima so beeinflussen, dass die Kinder ihren Mitschülern/innen helfen und sie unterstützen, auch wenn sie z.B. durch das Down- Syndrom extrem langsam arbeiten.“	.444	.531	.465	.594	2	.555
15 „Ich kann es schaffen, das Schulleben so zu gestalten, dass auch Kinder mit schwierigem Sozialverhalten darin einbezogen werden können.“	(.555)	.418	.494	.483	2	.488
11 „Auch bei größter Anstrengung wird es mir nicht gelingen, dass die Schüler die Leistungen und Beiträge von Kindern mit schwersten körperlichen Behinderungen wertschätzen.“ *	.323	-	(.507)	.450	4	.412

Anmerkungen. Alle Ladungen < .3 gelöscht. Grau unterlegte Felder markieren die Items der Kurzsкала KUS-HL. *zu rekodierende Items. Die Nummerierung der Items ist übernommen von Kopp (2009)

das Item 7: *„Ich kann Unterricht methodisch so gestalten, dass auch körperlich schwerstbehinderte Kinder ein für sie angemessenes Arbeitsangebot und einen passenden Arbeitsrhythmus finden können“.*

3. Im dritten Faktor finden sich indirekte Wirkungen auf die SuS; die Items beziehen sich auf die Selbstregulationskompetenz der Lehrkraft (siehe Tabelle 3). Vier Items finden sich bereits bei Kopp im Faktor „Inklusive Lehrerpersönlichkeit“ zusammen. Die inhaltliche Passung ist durch den gemeinsamen Bezug auf das Würdigen von SuS bzw. ihren Leistungen gegeben, auch wenn diese deviantes Verhalten zeigen. Auf diesen Aspekt bezieht sich auch Item 5, das im neuen Modell diesen Faktor ergänzt: *„Ich glaube, dass ich es auch beim besten Willen nicht schaffen werde, die Leistungen aller Kinder gleichermaßen zu würdigen und ihren Voraussetzungen entsprechend zu beurteilen, selbst wenn sie das Lernziel nicht erreicht haben“.* Die Bezeichnung „Inklusive Lehrerpersönlichkeit“ könnte auch für ein breiter angelegtes Konzept Verwendung finden, dennoch ist die Passung gegeben und der Begriff wird im Weiteren für diesen Faktor beibehalten.
4. Faktor 4 besteht – wie bereits benannt – aus drei Items, die sich auf das „Klassenklima“ beziehen, bzw. Anerkennungsprozesse zwischen SuS zum Inhalt haben. Item 11 *„Auch bei größter Anstrengung wird es mir nicht gelingen, dass die Schüler die Leistungen und Beiträge von Kindern mit schwersten körperlichen Behinderungen wertschätzen“* wurde zusätzlich aus inhaltlichen Gründen aufgenommen, auch wenn die Ladung auf Faktor 3 „Inklusive Lehrerpersönlichkeit“ etwas höher ausfällt (.450 zu .507).

Die Fragestellung 2 kann damit positiv beantwortet werden. Es kann ein alternatives und gut interpretierbares Modell beschrie-

ben werden, das auf der Grundlage einer Stichprobe von Studierenden verschiedener allgemeinbildenden Lehrämter gebildet wurde. Wie bereits dargestellt, achtet Kopp bei der Konstruktion der Items darauf verschiedene Differenzlinien in den Blick zu nehmen, da die Annahme bestand, dass diese sich auch auf Ebene der Faktorenstruktur zeigen würden. Entgegen den empirischen Ergebnissen von Kopp konnte durch die Unterscheidung der Faktoren 1 und 2 in der vorliegenden Studie diese Annahme teilweise bestätigt werden. Die Selbstwirksamkeit der Studierenden in Bezug auf SuS mit sonderpädagogischen Förderbedarfen bildet zwar einen korrelierten, aber dennoch eigenen Faktor, in Abgrenzung zur Selbstwirksamkeitserwartung in Bezug auf eine nicht weiter spezifizierte Leistungsheterogenität zusammen mit SuS mit migrationsbedingten sprachlichen Schwierigkeiten.

Konstruktion der Kurzsкала (KUS-HL)

Für die Konstruktion „Kurzsкала zur Selbstwirksamkeit im Umgang mit heterogenen Lerngruppen“ (KUS-HL) wurde aus Gründen der Testökonomie eine Obergrenze von 12 Items formuliert, wobei ein Minimum von drei Indikatoren pro Faktor aus methodischen Gründen zu berücksichtigen ist (vgl. Kline, 2015). Das favorisierte Modell mit vier Faktoren ermöglicht es beide Vorgaben umzusetzen. Weitere im Rahmen des Projekts ProfaLe entwickelte Kurzscales, z.B. zur Teamarbeit im inklusiven Unterricht, beinhalten ebenfalls 12 Items.

Die Auswahl der Items ist statistisch gelenkt und zielt auf eine möglichst hohe Ladung auf den jeweiligen Faktor. Ergänzend wurden die Trennschärfe der Items, Doppelloadungen und die von Mplus bereitgestellten Modifikationsindizes berücksichtigt. Als weitere Kriterien wurde darauf geachtet, dass alle Items inhaltlich unterschiedliche Facetten der latenten Variable repräsentieren und – wie auch in den anderen Kurz-

skalen – einige Items negativ formuliert sind. Alle für die Kurzskala ausgewählten Items sind in Tabelle 3 grau hinterlegt und jeweils als erstes aufgeführt.

Konfirmatorische Überprüfung alternativer Messmodelle

Analysemethode CFA für alternative Modelle

Die konfirmatorische Überprüfung der neuen Faktorenstruktur erfolgt anhand der Stichprobe 1 ($n = 565$), die auch für die CFA des Kopp-Modells verwendet wurde. Das methodische Vorgehen ist ebenfalls das gleiche. Wie bereits in Abschnitt „Ergebnisse EFA“ dargestellt, weisen die EFA anhand der Stichprobe 2 ($n = 384$) auf eine zwei- oder eine vierfaktorielle Lösung hin. Die genestete Relation dieser beiden Alternativen ermöglicht den direkten Vergleich mittels eines skalierten Chi-Quadrat-Differenztests. Die konfirmatorische Überprüfung der Kurzskala erfolgt an der Stichprobe 3 ($n = 456$). In einem Step-Up-Ansatz (vgl. Schwab & Helm, 2015; Kleinke, Schlüter & Christ, 2017) wird ergänzend die Messinvarianz geprüft, d.h. der Frage nachgegangen, ob die Skalen für verschiedene Lehrämter das gleiche Konstrukt in derselben Weise messen.

Ergebnisse CFA der alternativen Modelle

Zunächst kann berichtet werden, dass die Lösung mit zwei Faktoren einen minimal schlechteren Fit als das Kopp-Modell mit korrelierten Faktoren aufweist (siehe Tabelle 4). Die favorisierte Lösung mit vier Faktoren weist hingegen einen deutlich besseren Fit auf. Die Overall-Fit-Indizes *RMSEA* (.041-.051; guter Fit: *RMSEA* < .05) und *SRMR* (.05; guter Fit: *SRMR* < .05; vgl. Schermelleh-Engel et al., 2003) liegen für dieses Modell knapp im bzw. an der Grenze zu einem guten Bereich. Bei den inkrementellen Indizes wären nach großzügigeren Grenzwerten der *CFI* (.914) und der *TLI* (.903) noch akzeptabel. Nach den etablierten strengeren Grenzwerten wird ein akzeptabler Fit vom *CFI* und vom *TLI* (*CFI/TLI* > .95) verfehlt (vgl. Kline, 2015; Schermelleh-Engel et al., 2003). Ein signifikanter Chi-Quadrat-Differenztest ($TRd = 109.67$; $\Delta df = 5$; $p < .001$) unterstreicht die Vorteile des komplexeren vierfaktoriellen Modells (vgl. Satorra & Bentler, 2001).

Wie erwartet verbessert sich der Fit für die Kurzskala, vermutlich insbesondere durch das Eliminieren mehrerer Items mit Doppel- bzw. Mehrfachladungen. Der *CFI* steigt auf einen akzeptablen Wert von .954. Der *TLI* gewinnt ebenfalls deutlich (.903 > .937) und nähert sich einem akzeptablen Fit an (siehe Tabelle 4).

Die Prüfung der hierarchischen Messinvarianzen für das neue Modell mit 24 Items

Tabelle 4: Anpassungsgüte der alternativen Modelle. Stichprobe 1 ($n = 565$) und 3 ($n = 456$) für die Kurzskala

Stichprobe	Faktor Modell	Items	χ^2	df	χ^2/df	$\chi^2 dif$ $p <$	CFI	TLI	RMSEA (90%CI)	SRMR
1	Kopp (oblique)	24	662.37	246	2.70	-	.881	.867	.049-.059	.053
1	2 Fak. neu	24	671.91	251	2.68	-	.880	.868	.049-.059	.055
1	4 Fak. neu	24	548.13	246	2.23	.001	.914	.903	.041-.051	.050
3	4 Fak. KUS-HL	12	102.23	48	2.13		.954	.937	.036-.063	.041

Anmerkungen. CFI: Comparative Fit Index; TLI: Tucker-Lewis Index; RMSEA: Root Mean Square of Approximation; 90% Konfidenzintervall für RMSEA-Schätzer; SRMR: Standardized Mean Square Residual. Signifikanzen im skalierten Chi-Quadrat-Differenztest berechnet nach Satorra und Bentler (2001).

Tabelle 5: CFA für getrennte Samples (LaGym, LaPriSe, LaSopäd) und 24 Items. Stichprobe 1 (n = 565)

Modell Stichprobe	Items	χ^2	df	χ^2/df	CFI	TLI	RMSEA (90%CI)	SRMR
Gesamt – 4 Faktoren	24	548.13	246	2.23	.914	.903	.041-.051	.050
LaGym – 4 Faktoren	24	399.12	246	1.62	.828	.807	.047-.068	.074
LAPriSe – 4 Faktoren	24	465.95	246	1.89	.872	.856	.049-.064	.064
LaSopäd – 4 Faktoren	24	340.88	246	1.39	.819	.797	.045-.078	.083

zeigt bereits im ersten Schritt des Step-Up-Ansatzes deutliche Probleme in der Passung für die drei Lehrämter der Stichprobe 1 auf. Alle Fit-Indizes fallen für die Berechnung pro Lehramtsgruppe erheblich ab (vgl. Tabelle 5). Somit kann nicht davon ausgegangen werden, dass die Items gleichermaßen für die verschiedenen Lehrämter den Faktoren zuzuordnen sind und sich damit die Langform mit 24 Items für jede weitere Verwendung disqualifiziert.

Stichprobe 3, anhand derer die Kurzskala getestet wird, enthält Daten nur für die Lehrämter PriSe und Sopäd. Auch in diesem Fall sinkt für die Gruppe Sopäd der Modellfit in der separaten Berechnung erheblich, sodass kein akzeptabler Modellfit mehr erreicht wird (vgl. Tab. 6). Die Modifikationsindizes verweisen auf Probleme in Faktor 3 „Inklusive Lehrerpersönlichkeit“. Eine erneute Berechnung des Modellfits für beide Studierendengruppen, unter dem Aus-

schluss von Faktor 3, zeigt einen guten Modellfit für beide Gruppen auf.

Der Modellfit für das Baseline-Modell ohne zusätzliche Restriktionen für 9 Items (3 Faktoren) ist durchgehend akzeptabel (vgl. Tabelle 6). Auch die Gleichheitsrestriktionen auf den Faktorladungen im Fall der metrischen Invarianz verschlechtern die Modellpassung nicht. Der ergänzende skalierte χ^2 -Test bestätigt die Annahme der metrischen Invarianz durch ein deutlich nicht signifikantes Ergebnis ($p = .30$). Die Maßeinheit der Skala ist damit in beiden Stichproben gleich. Auch die Gleichsetzung der Intercepts und die Freisetzung der Mittelwerte der latenten Variablen zur Überprüfung der skalaren Invarianz verschlechtern den Modellfit nur unbedeutend. Der skalierte χ^2 -Test ist im Vergleich mit dem metrischen Invarianzmodell erneut nicht signifikant ($p = .127$), somit können auch Mittel-

Tabelle 6: Ergebnisse der Überprüfung der Messinvarianz für die Kurzskala KUS-HL mit 12 und alternativ 9 Items. Stichprobe 3 (n = 456)

Modell Stichprobe	Items	χ^2	df	χ^2/df	χ^2/dif p	CFI	TLI	RMSEA (90%CI)	SRMR
Gesamt	12	102.23	48	2.13	-	.954	0.937	.036-.063	.041
Sopäd – 4 Faktoren	12	90.15	48	1.88	-	.869	.820	.049-.095	.068
Sopäd – 3 Faktoren	9	29.47	24	1,23	-	.976	.964	.000-.76	.047
PriSe – 4 Faktoren	12	90.97	48	1,90	-	.934	.910	.042-.080	.049
PriSe – 3 Faktoren	9	40.24	24	1,67	-	.967	.950	.021-.081	.040
Baseline-Modell	9	68.48	48	1,42	-	.970	.955	.016-.069	.043
Metrische Invarianz	9	75.78	54	1,40	.30	.968	.958	.016-.067	.061
Skalare Invarianz	9	85.41	60	1,42	.13	.963	.956	.020-.066	.061

wertvergleiche zwischen den beiden Lehrämtern durchgeführt werden.

Ergänzend wurden alle vier Subskalen mittels Mann-Whitney-Test auf Gruppenunterschiede überprüft. Die Selbstwirksamkeit fällt in den Skalen „Adaptive Unterrichtsgestaltung in heterogenen Lerngruppen“ ($M_{Sopäd} = 4.39$, $M_{PriSe} = 3.70$, $Z = -6,69$, $p < .001$), „Inklusive Lehrerpersönlichkeit“ ($M_{Sopäd} = 5.21$, $M_{PriSe} = 4.87$, $Z = -3,76$, $p < .001$) und „Klassenklima“ ($M_{Sopäd} = 5.12$, $M_{PriSe} = 4.75$, $Z = -4,11$, $p < .001$) signifikant höher für die Studierenden Sopäd aus. Die Aussagekraft ist hierbei durch die fehlende Messinvarianz zwischen den Gruppen für die Subskala „Inklusive Lehrerpersönlichkeit“ eingeschränkt. Es besteht kein signifikanter Unterschied zwischen den Gruppen in der Subskala 2 „Adaptive Unterrichtsgestaltung für SuS mit erhöhtem sonderpädagogischen Förderbedarf“ ($M_{Sopäd} = 4.06$, $M_{PriSe} = 3,91$, $Z = -1,92$, $p = .054$), die – so wäre anzunehmen – insbesondere sonderpädagogische Aufgabenbereiche abbildet.

Diskussion der Ergebnisse

Selbstwirksamkeitserwartungen von Lehrkräften stellen eine Gelingensbedingung für die erfolgreiche Gestaltung von inklusiven Schul- und Unterrichtsettings dar. Entsprechend ist deren zuverlässige und valide Messung ein wichtiges Desiderat empirischer Forschung. Das von Kopp (2009) vorgelegte Instrument war ein wichtiger Ausgangspunkt zur Erforschung dieser Thematik im deutschsprachigen Raum. Die zwischenzeitliche Etablierung und Ausgestaltung eines wissenschaftlichen Diskurses wurde u.a. durch eine Reihe weiterer Forschungsprojekte mit explizitem Bezug zu ihren Vorarbeiten vollzogen. Die vorliegende Untersuchung hatte zum Ziel, die bisher angenommene Differenzierung in verschiedene Facetten von Selbstwirksamkeit nach Kopp (2009) zu überprüfen, da diese bisher nur explorativ und nur mit einer Stichprobe von

Studierenden mit dem Lehramt Grundschule überprüft wurde. Hierbei wurden für die explorativen und konfirmativen Analysen jeweils unterschiedliche Stichproben verwendet bzw. eine dritte für die konfirmative Überprüfung der Kurzska.

Im Ergebnis zeigte sich, dass das Modell von Kopp mit unkorrelierten Faktoren nicht zur Stichprobe 1 aus Studierenden der Lehrämter Sonderpädagogik, Primar- und Sekundarstufe sowie Gymnasium passt (Fragestellung 1). Ein Modell mit gleicher Zuordnung der Items, aber mit korrelierten Faktoren bietet hingegen einen akzeptablen bis mittelmäßigen Fit. Es konnte außerdem auf Grundlage einer zweiten Stichprobe mit Anteilen von Studierenden der verschiedenen allgemeinbildenden Lehrämter und mit Hilfe explorativer Faktorenanalysen ein alternatives und gut interpretierbares Modell extrahiert werden (Fragestellung 2). Inhaltlich zeigt sich die größte Diskrepanz zum Modell von Kopp darin, dass die neue vierfaktorielle Struktur Selbstwirksamkeit im Umgang mit einer nicht weiter spezifizierten Leistungsheterogenität (Faktor 1) und einer solchen mit einem Fokus auf Behinderung unterscheidet (Faktor 2).

Ob die gefundene Faktorenstruktur eine hohe Passung zur (für die konfirmativen Analysen verwendeten) Stichprobe 1 aufweist, kann weder pauschal mit „ja“ noch mit „nein“ beantwortet werden (Fragestellung 3). Zunächst ist festzuhalten, dass das gut interpretierbare neue vierfaktorielle Modell einen deutlich besseren Fit aufwies, allerdings nicht in allen Indizes akzeptable Werte erzielte. Die an der dritten Stichprobe überprüfte Kurzska zeigte hingegen eine akzeptable bis gute Passung und verfehlte nur im TLI knapp einen akzeptablen Fit.

Ergänzend zeigten sich in den Faktoren 1, 3 und 4 signifikant höhere Mittelwerte für die Studierenden der Sonderpädagogik im Vergleich zu den Studierenden der Primar und/oder Sek.1. Für die vorliegende Studie ist weniger die Richtung dieser Differenz interessant als der implizit geführte Nachweis

einer grundsätzlichen Sensitivität der Kurzskala für solche Gruppenunterschiede. Dabei erreichten alle vier Skalen das Niveau metrischer Invarianz, was Vergleiche der Strukturbeziehungen zwischen den Gruppen legitimiert. Die Modifikationsindizes zeigten allerdings Passungsprobleme für den Faktor 3 „Inklusive Lehrerpersönlichkeit“ auf und das Niveau der skalaren Messinvarianz wurde erst durch den Ausschluss dieses dritten Faktors erreicht. Es ist somit nicht sichergestellt, dass gleiche Konstruktwerte in Faktor 3 auch die gleichen Indikatorwerte bedingen, sodass Mittelwertunterschiede nur unter diesem Vorbehalt interpretiert werden können.

Auf Grundlage der getätigten Analysen kann die Verwendung der Skalen von Kopp in bisherigen Forschungsvorhaben mit Einbezug verschiedener Gruppen von Studierenden kritisch beurteilt werden. Die neue Faktorenstruktur und die neue Kurzskala KUS-HL bieten zukünftig eine Alternative und könnten auch zur Reanalyse bereits durchgeführter Projekte dienen.

Die Sensitivität der Skala für Veränderungen der Selbstwirksamkeit in Längsschnitterhebungen gilt es in zukünftigen Forschungsprojekten zu überprüfen. Außerdem sollten die Ursachen für die nicht gegebene Messinvarianz in Faktor 3 „Inklusive Lehrerpersönlichkeit“ weiter analysiert werden.

Als Hauptkritik kann die fehlende Dokumentation der Itemkonstruktion und der dahinterstehenden Systematik herausgestellt werden. Eine systematische Konstruktion von Items, z.B. entsprechend einem facettentheoretischen Ansatz (vgl. Bilsky & Cairns, 2009) oder Vorstudien mit einem deutlich größeren Itempool und schrittweiser Selektion der Items, wäre wünschenswert. Vielfältige Hinweise für eine solche Neukonstruktion einer Selbstwirksamkeitsskala mit dem Bezug auf Unterricht in inklusiven Settings liefert die vorliegende Studie.

Literaturverzeichnis

- Abele, A. E. & Candova, A. (2007). Prädiktoren des Belastungserlebens im Lehrerberuf. *Zeitschrift für Pädagogische Psychologie*, 21(2), 107–118.
- Baltes-Götz, B. (2013). *Behandlung fehlender Werte in SPSS und Amos*. Verfügbar unter: <https://www.uni-trier.de/fileadmin/urt/doku/bfw/bfw.pdf> [30.04.2018].
- Bandura, A. (1997). *Self-efficacy: the exercise of control*. New York: Freeman.
- Bilsky, W. & Cairns, D. (2009). Facettentheorie. In H. Holling (Hrsg.), *Grundlagen und statistische Methoden der Evaluationsforschung* (S. 135-162). Göttingen: Hogrefe.
- Blömeke, S., Gustafsson, J. E. & Shavelson, R. (2015). Beyond dichotomies: competence viewed as a continuum. *Zeitschrift für Pädagogik*, 223(1), 3-13.
- Boer, A. de, Pijl, S. J. & Minnaert, A. (2011). Regular primary schoolteachers' attitudes towards inclusive education: A review of the literature. *International Journal of Inclusive Education*, 15 (3), 331-353.
- Bosse, S., Henke, T., Jäntsch, C., Lambrecht, J., Maaz, K., Vock, M., Schründer-Lenzen, A. & Spörer, N. (2016). Zum Zusammenhang von Einstellung und Selbstwirksamkeit von inklusiv arbeitenden Grundschullehrkräften. In K. Liebers, B. Landwehr, S. Reinhold, S. Riegler & R. Schmidt (Hrsg.), *Facetten grundschulpädagogischer und -didaktischer Forschung* (S. 99-104). Wiesbaden: Springer.
- Bosse, S. & Spörer, N. (2014). Erfassung der Einstellung und der Selbstwirksamkeit von Lehramtsstudierenden zum inklusiven Unterricht. *Empirische Sonderpädagogik*, 6 (4), 279-299.
- Cloerkes, G. (2007). *Soziologie der Behinderten: eine Einführung*. Heidelberg: Winter.
- Costello, A. B. & Osborne, J. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, 10(7). Available online: <http://pareonline.net/getvn.asp?v=10&n=7>.

- Depaepe, F. & König, J. (2018). General pedagogical knowledge, self-efficacy and instructional practice: Disentangling their relationship in pre-service teacher education. *Teaching and Teacher Education*, 44 (2), 142-155.
- Dessemontet, R. S., Morin, D., & Crocker, A. G. (2014). Exploring the Relations between In-service Training, Prior Contacts and Teachers' Attitudes towards Persons with Intellectual Disability. *International Journal of Disability, Development and Education*, 61 (1), 16-26.
- Forlin, C., Loreman, T., & Sharma, U. (2014). A system-wide professional learning approach about inclusion for teachers in Hong Kong. *Asia-Pacific Journal of Teacher Education*, 42(3), 247-260.
- Gebauer, M. M. (2013). *Determinanten der Selbstwirksamkeitsüberzeugung von Lehrenden*. Wiesbaden: Springer.
- Grosche, M. (2015). Was ist Inklusion? Ein Diskussions- und Positionsartikel zur Definition von Inklusion aus Sicht der empirischen Bildungsforschung. In P. Kuhl et al. (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 17-39). Wiesbaden: Verlag für Sozialwissenschaften.
- Heinrich, M., Urban, M. & Werning, R. (2013). Grundlagen, Handlungsstrategien und Forschungsperspektiven für die Ausbildung und Professionalisierung von Fachkräften für inklusive Schulen. In: H. Döbert & H. Weishaupt (Hrsg.): *Inklusive Bildung professionell gestalten. Situationsanalyse und Handlungsempfehlungen*. Münster: Waxmann.
- Hellmich, F. & Görel, G. (2014). Erklärungsfaktoren für Einstellungen von Lehrerinnen und Lehrern zum inklusiven Unterricht in der Grundschule. *Zeitschrift für Bildungsforschung*, 4 (3), 227-240.
- Hellmich, F., Görel, G. & Schwab, S. (2016). Einstellungen und Motivation von Lehramtsstudentinnen und -studenten in Bezug auf den inklusiven Unterricht in der Grundschule – Ein Vergleich zwischen Deutschland und Österreich. *Empirische Sonderpädagogik*, 8 (1), 67-85.
- Kaiser, H. F. & Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*, 34 (1), 111-117.
- Kleinke, K., Schlüter, E. & Christ, O. (2017). *Strukturgleichungsmodelle mit Mplus – Eine praktische Einführung*. Oldenburg: De Gruyter.
- Kline R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: Guilford.
- Knigge, M. & Rotter, C. (2014). *EiLInk – Einstellungen von Lehramtsstudierenden zu Inklusion. Skalenhandbuch* – erstellt unter Mitarbeit von K. Thomas & L. Wu. – nicht publiziert.
- Knigge, M. & Rotter, C. (2015). Unterrichtsplanungen bei Lehramtsstudierenden im Falle der Wahrnehmung von vermeintlich „besonderen“ Schülerinnen und Schülern und ihr Zusammenhang mit Selbstwirksamkeitsüberzeugungen und Einstellungen in Bezug zu Inklusion – beispielhafte Mixed-Method-Analysen aus der EiLInk-Studie. *Empirische Sonderpädagogik*, 7 (3), 223-240.
- Kopmann, H. & Zeinz, H. (2016). Einstellungen zur Inklusion in der Grundschule und pädagogisch-didaktische Orientierungen. In K. Liebers, B. Landwehr, S. Reinhold, S. Riegler, R. Schmidt (Hrsg.), *Facetten grundschulpädagogischer und grundschuldidaktischer Forschung* (S.75-80). Wiesbaden: Springer VS.
- Kopp, B. (2009). Inklusive Überzeugung und Selbstwirksamkeit im Umgang mit Heterogenität. Wie denken Studierende des Lehramts für Grundschulen? *Empirische Sonderpädagogik*, 1 (1), 5-25.
- Krause, A., Dorsewagen, C., & Baeriswyl, S. (2013). Zur Arbeitssituation von Lehrerinnen und Lehrern: Ein Einstieg in die Lehrerbelastungs- und -gesundheitsforschung. In M. Rothland (Hrsg.), *Belastung und Beanspruchung im Lehrerberuf* (S.61-80). Wiesbaden: Springer.

- Little, T. D. (2013). *Longitudinal structural equation modeling*. New York, NY: Guilford Press.
- Loreman, T., Sharma, U. & Forlin, C. (2013). Do Pre-service Teachers Feel Ready to Teach in Inclusive Classrooms?: A Four Country Study of Teaching Self-efficacy. *Australian Journal of Teacher Education*, 38 (1), 27-44.
- Malinen, O.-P., Savolainen, H., Engelbrecht, P., Xu, J., Nel, M., & Nel, N. (2013). Exploring teacher self-efficacy for inclusive practices in three diverse countries. *Teaching and Teacher Education*, 33, 34-44.
- Melzer, C., Hillenbrand, C., Sprenger, D. & Hennemann, T. (2015). Aufgaben von Lehrkräften in inklusiven Bildungssystemen – Review internationaler Studien. *Erziehungswissenschaft*, 26 (51), 61- 80.
- Moosbrugger, H., & Schermelleh-Engel, K. (2012). Exploratorische (EFA) und Konfirmatorische Faktorenanalyse (KFA). In H. Moosbrugger & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 326-343). Berlin: Springer.
- Park, M.-H., Dimitrov, D. M., Das, A. & Gichuru, M. (2016). The teacher efficacy for inclusive practices (TEIP) scale: Dimensionality and factor structure. *Journal of Research in Special Educational Needs*, 16 (1), 2-12.
- Richter, D., Engelbert, M., Weirich, S., & Pant, H.A. (2013). Differentielle Teilnahme an Lehrerfortbildungen und deren Zusammenhang mit professionsbezogenen Merkmalen von Lehrkräften. *Zeitschrift für Pädagogische Psychologie*, 27 (3), 193-207.
- Satorra, A., & P.M. Bentler (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66, 507-514.
- Sharma, U. & Nuttal, A. (2016). The impact of training on pre-service teacher attitudes, concerns, and efficacy towards inclusion. *Asia-Pacific Journal of Teacher Education*, 44 (2), 142-155.
- 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.
- Schmitz, G. S. & Schwarzer, R. (2000). Selbstwirksamkeitserwartung von Lehrern: Längsschnittbefunde mit einem neuen Instrument. *Zeitschrift für Pädagogische Psychologie*, 14 (1), 12-25.
- Schmitz, G. S. & Schwarzer, R. (2002). Individuelle und kollektive Selbstwirksamkeitserwartung von Lehrern. *Zeitschrift für Pädagogik*, 44, Beiheft: Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen, 192-214.
- Schwab, S. & Helm, C. (2015). Überprüfung von Messinvarianz mittels CFA und DIF-Analysen. *Empirische Sonderpädagogik*, 3, 175-193.
- Schwarzer, R. & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor: NFER-NELSON.
- Sharma, U. & Nuttal, A. (2015). The impact of training on pre-service teacher attitudes, concerns, and efficacy towards inclusion. *Asia-Pacific Journal of Teacher Education*, 44 (2), 142-155.
- Urban, Dieter; Mayerl, Jochen (2014). *Strukturgleichungsmodellierung. Ein Ratgeber für die Praxis*. Wiesbaden: Springer VS.
- Watkins, M. W. (2006). Determining Parallel Analysis Criteria. *Journal of Modern Applied Statistical Methods*, 5 (2), 344-346.
- Watkins, M.W. (2008). *Monte Carlo PCA for Parallel Analysis. Version 2.3*. (Computer Software). Verfügbar unter: <http://edpsychassociates.com/Watkins3.html>
- Weisel, A. & Dror, O. (2006). School climate, sense of efficacy and Israeli teachers' attitudes toward inclusion of students with special needs. *Education, Citizenship and Social Justice*, 1 (2), 157-174.
- Wentura, D. & Pospeschill, M. (2015). *Multivariate Datenanalyse – Eine kompakte Einführung*. Wiesbaden: Springer.

Werning, R. (2014). Stichwort: Schulische Inklusion. *Z Erziehungswiss.*, 17, 601-623.

Woodcock, S., Hemmings, B. & Kay, R. (2012). Does Study of an Inclusive Education Subject Influence Pre-Service teachers' Concerns and Self-Efficacy about Inclusion? *Australian Journal of Teacher Education*, 37 (6), 1-11.

Wudy, D.-T. & Jerusalem, M. (2011). Die Entwicklung von Selbstwirksamkeit und Belastungserleben bei Lehrkräften. *Psychologie in Erziehung und Unterricht*, 58, 254-267.

Dr. Steffen Siegemund

Julius-Maximilians-Universität Würzburg
Lehrstuhl für Pädagogik bei
Geistiger Behinderung
Wittelsbacher Platz 1
97074 Würzburg, Germany
steffen.siegemund-johannsen@uni-
wuerzburg.de

Erstmalig eingereicht: 18.03.2019

Überarbeitung eingereicht: 06.08.2019

Angenommen: 06.08.2019