

Untersuchung der Einstellung zu Inklusion: Validierung und faktorenanalytische Überprüfung einer Kurzversion der PREIS-Skala

Gunnar Bruns  · Timo Lücke  · Cornelia Gresch  ·
Michael Grosche 

Eingegangen: 14. Dezember 2022 / Überarbeitet: 8. Mai 2023 / Angenommen: 10. Mai 2023 / Online
publiziert: 21. Juni 2023
© Der/die Autor(en) 2023

Zusammenfassung Der Zusammenhang von Einstellungen mit Verhaltensweisen gilt gemäß der Theorie des geplanten Verhaltens im Allgemeinen als gut belegt. Um untersuchen zu können, ob eine positive Einstellung zu einem inklusiven Schulsystem mit Verhaltensweisen zusammenhängt, welche die Umsetzung von Inklusion begünstigen, wurde eine Professionsunabhängige Einstellungsskala zum Inklusiven Schulsystem (PREIS) mit 14 Items entwickelt. In diesem Beitrag wird für die Kurzskala (PREIS-K) mit kurzer Instruktion und 6 Items geprüft, ob die eindimensionale Faktorenstruktur, die Messäquivalenz zwischen Personengruppen und die konvergente Validität weiterhin gegeben sind. $N=1038$ Lehrkräfte und $N=1438$ Eltern beantworteten im Rahmen einer nationalen Vergleichsstudie zur Inklusion in der Sekundarstufe die PREIS-K. Ein Bifaktor-Modell mit Berücksichtigung eines Methodenfaktors für positiv und negativ formulierte Items passt besser zu den Daten als ein Einfaktor-Modell. Die Messinvarianz zwischen Lehrkräften und Eltern ist auf skalarer Ebene gegeben, innerhalb der Lehrkräfte (zwischen sonderpädagogischen und allgemeinen Lehrkräften) kann sogar von residualer Messinvarianz ausgegangen werden. Positive hypothesenkonforme Korrelationen mit konvergenten Variablen

✉ Dr. Gunnar Bruns · Prof. Dr. Michael Grosche
Bergische Universität Wuppertal, Wuppertal, Deutschland
E-Mail: gunnar.bruns@uni-wuppertal.de

Prof. Dr. Michael Grosche
E-Mail: michael.grosche@uni-wuppertal.de

Prof. Dr. Timo Lücke
Karl-Franzens-Universität Graz, Graz, Österreich
E-Mail: timo.luecke@uni-graz.at

Forschungszentrum Inklusive Bildung, Graz, Österreich

Dr. Cornelia Gresch
Humboldt-Universität zu Berlin, Berlin, Deutschland
E-Mail: cornelia.gresch@hu-berlin.de

sind nachweisbar. Insgesamt ermöglicht die PREIS-K eine valide und ökonomische Überprüfung der Einstellungen zu einem inklusiven Schulsystem. Vergleiche der Einstellungen von Personengruppen sind mit der PREIS-K möglich. Aufgrund ihrer Kürze sowie den bereits durchlaufenen ministeriellen Genehmigungsverfahren eignet sich die PREIS-K auch für groß angelegte Vergleichs- und Längsschnittstudien.

Schlüsselwörter Einstellungen · Inklusion · Lehrkräfte · Eltern · Faktorenstruktur

Measuring attitudes towards inclusive education: testing the validity and factor structure of a short version of the PREIS-scale

Abstract Attitudes are generally closely connected with behavior, according to the theory of planned behavior. In order to investigate the relation between a positive attitude towards inclusive education and those behaviors that are likely to promote inclusion, a scale ‘Professionsunabhängige Einstellungsskala zum Inklusiven Schulsystem (PREIS)’ [Attitudes Towards Inclusive Education in the Population] was developed, consisting of 14 items. In the current study, we test if a shortened scale (PREIS-K) with a reduced instruction and a selection of 6 items still satisfies the one-dimensional factor structure, measurement equivalence between different groups, and convergent validity. A total of $N=1038$ teachers and $N=1438$ parents, who participated in the INSIDE-study, answered the PREIS-K. A bifactor model that included a method factor yielded a better fit than a one-dimensional model. Between teachers and parents, the measure is invariant on the scalar level, and within teachers (regular and special education staff) on the residual level. Positive correlations with convergent variables can be shown. Overall, the PREIS-K allows a valid and parsimonious measure of attitudes towards inclusive education. Comparing attitudes across different groups is possible. As the PREIS-K is very brief and has already been approved by the ministry of education, it is also suitable for large scale assessments.

Keywords Attitudes · Inclusion · Teachers · Parents · Factor structure

1 Einleitung

Lehrkräfte mit positiven Einstellungen zeigen häufiger günstiges Verhalten im Sinne erfolgreicher Umsetzung schulischer Inklusion (u. a. Mosen et al. 2014; Wilson et al. 2016). Daher werden positive Einstellungen als hilfreiche Ressource für das Gelingen von schulischer Inklusion aufgefasst (z. B. Avramidis und Norwich 2002; de Boer et al. 2011). Um diesen theoretisch gut begründeten Zusammenhang empirisch abzusichern, bedarf es geeigneter Messinstrumente, die die Einstellungen zur schulischen Inklusion valide erfassen, was jedoch nicht bei allen derzeit verfügbaren Verfahren der Fall ist (Lüke und Grosche 2018). Die Professionsunabhängige Einstellungsskala zum Inklusiven Schulsystem (PREIS) wurde in einer Reihe von unabhängigen Studien empirisch überprüft und validiert (Lüke und Grosche 2018; Röhms et al. 2018; Schulze et al. 2019), und gilt damit als eine der am sorgfältigsten

entwickelten, dokumentierten und validierten Skalen zur Messung der Einstellung zu Inklusion (Kielblock und Woodcock 2023). Jedoch handelt es sich bei der PREIS um ein recht umfangreiches Instrument. Um für breiter angelegte Studien ein ökonomischeres Instrument zur Verfügung zu stellen, kürzten Lüke und Grosche (2020) die Instruktion und reduzierten die Itemanzahl deutlich. Die Messgüte dieser Kurzskaala PREIS-K (Faktorenstruktur, Messinvarianz und Konstruktvalidität) wird in der vorliegenden Studie geprüft. Wenn sich die PREIS-K Skala auch in ihrer gekürzten Fassung als robust erweist, stellt dies eine gute empirische Begründung für deren Einsatz dar – auch in groß angelegten Vergleichs- und Längsschnittstudien, die auf besonders ökonomische Skalen angewiesen sind.

1.1 Theoretische Grundlagen der PREIS und PREIS-K

Der Begriff der schulischen Inklusion ist diffus und äußerst vielschichtig. Folglich unterscheiden sich auch die Einstellungen, je nachdem zu welcher Facette von Inklusion Personen befragt werden. Daher sollte explizit definiert werden, welche Definition von Inklusion einem Instrument zugrunde liegt bzw. worauf genau die befragten Personen ihre Einstellungen beziehen sollen. Die wenigsten Instrumente definieren das Einstellungsobjekt explizit. Die gemessenen Einstellungen sind daher möglicherweise nur schlecht miteinander vergleichbar.

Göransson und Nilholm (2014) unterscheiden vier Definitionsebenen von schulischer Inklusion, die allesamt wichtige Facetten des Inklusionsdiskurses widerspiegeln und anhand ihrer Komplexität hierarchisch geordnet sind. Die Platzierungsdefinition (A) meint die gemeinsame Beschulung von Schüler*innen mit und ohne sonderpädagogische Förderbedarfe (SPF). Die spezifische Individualisierungsdefinition (B) meint die Ermöglichung möglichst hoher akademischer Kompetenzen und sozialer Erlebnisse bei Schüler*innen mit SPF. Die allgemeine Individualisierungsdefinition (C) zielt zusätzlich auf die akademischen Kompetenzen und sozialen Erlebnisse *aller* Schüler*innen ab, unabhängig von SPF. Die Gemeinschaftsdefinition (D) beschreibt eine besondere Lerngemeinschaft, die beispielsweise demokratisch, selbstbestimmend, partizipativ und anerkennend sowie nicht-exklusiv ist. Die PREIS bzw. PREIS-K nutzt die dritte Definition: „In einem inklusiven Schulsystem besuchen alle Kinder eine gemeinsame Schule. Die Schulen fühlen sich verantwortlich für die Förderung und Unterstützung aller Kinder – unabhängig von ihren Lernvoraussetzungen, Erfahrungen und Bedürfnissen.“ (Lüke und Grosche 2020, S. 1). Bei dem zu bewertenden Einstellungsgegenstand geht es, mit Blick auf die gesamtgesellschaftliche breite Zielgruppe der Skala, vorrangig um ein auf Partizipation ausgerichtetes inklusives Schulsystem als Ganzes, und weniger um spezifische Aspekte von (adaptiver) Unterrichtsgestaltung oder konkretem Lehrerhandeln.

Wir verwenden den Begriff der Einstellungen in Anlehnung an Eagly und Chaiken (2007) für die psychologische Tendenz, einen bestimmten Einstellungsgegenstand (hier das inklusive Schulsystem) auf der Basis einer *summativen* Bewertung kognitiver (z.B. Wissen über Vor- und Nachteile eines inklusiven Schulsystems), affektiver (z.B. Emotionen in Bezug auf Menschen mit bestimmten Behinderungen) und behavioraler Informationen (z.B. eigenes bisheriges Verhalten oder bestimmte Werthaltungen) als eher positiv oder negativ zu beurteilen (vgl. Schwarz und Bohner

2001; Eagly und Chaiken 2007). Wir gehen dabei davon aus, dass diese summative Evaluation zwar auf kognitiven, affektiven und behavioralen Erfahrungswerten basiert, es aber „kaum eine Rechtfertigung für die Behauptung [gibt], dass diese Erfahrungen drei trennbare und allgegenwärtige Komponenten von Bewertungstendenzen“ (Eagly und Chaiken 2007, S. 590) in Form von Subskalen ergäben. Auch bei der faktorenanalytischen Prüfung der Fragebögen im Themenfeld der Einstellungen zum inklusiven Schulsystem zeigt sich eine Aufteilung auf drei entsprechende Subskalen in der Regel nicht.

Zur Messung von Einstellungen zu schulischer Inklusion steht eine Reihe von Instrumenten zur Verfügung, die bei Lüke und Grosche (2018) einer ausführlichen Kritik unterzogen werden. Zusammengefasst lassen sich zwei wesentliche Kritikpunkte ausmachen: a) häufig kommen Instrumente zum Einsatz, deren inhaltliche und methodische Fundierung nicht immer ausreichend gegeben ist (so lassen sich z. B. getrennte kognitive, affektive und behaviorale Komponenten faktorenanalytisch meist nicht nachweisen) und b) richten sich viele Instrumente spezifisch an Lehrkräfte, was bei Personen ohne pädagogische Vorerfahrungen dazu führt, dass eine vergleichbar valide Messung mit diesen Instrumenten nicht zu erwarten ist. Daher wurde die PREIS entwickelt, die aus einer ausführlichen Instruktion und 14 positiv und negativ formulierten Items besteht (z. B. „Ich finde, ein inklusives Schulsystem wäre gerechter als das jetzige Schulsystem“). Die Beantwortung erfolgt auf einer fünfstufigen Likert-Skala mit semantischen Ankern an den Endpunkten (0= „stimme nicht zu“ und 4= „stimme zu“). Zur empirischen Absicherung der PREIS-Skala wurden mehrere unabhängige Studien durchgeführt, die eine inhaltlich und faktoriell valide sowie konsistente und reliable Messung mehrfach repliziert haben (z. B. Lüke und Grosche 2018; Röhm et al. 2018; Schulze et al. 2019). In Übereinstimmung mit unseren Annahmen über die Struktur von Einstellungen werden in der PREIS in den Itemtexten zwar kognitive, affektive und behaviorale Einstellungsinhalte berücksichtigt, die Struktur ist aber theoretisch und auch empirisch eindimensional im Sinne einer summativen Evaluation des Einstellungsgegenstandes.

Die PREIS stellt folglich eine empirisch bewährte Möglichkeit zur Erfassung von Einstellungen zu einem inklusiven Schulsystem dar. Für den Einsatz in groß angelegten Vergleichs- oder Längsschnittstudien der Bildungsforschung ist das Instrument jedoch aufgrund der Länge der Instruktion und der Itemanzahl nicht ökonomisch genug. In solchen Studien werden in der Regel übergeordnete Ziele verfolgt, wie z. B. die Evaluation der Umsetzung von Bildungsstandards in den Bundesländern (IQB Bildungstrend), internationale Vergleiche schulischer Kompetenzen (z. B. PISA) oder die Untersuchung von Bildungsverläufen und Kompetenzentwicklung (Nationales Bildungspanel, NEPS). Im Mittelpunkt dieser Studien steht meist die umfassende Messung schulischer Kompetenzen. Ergänzend erhalten Schüler*innen, Lehrkräfte und Schulleitungen Fragebögen, in denen Hintergrundinformationen erhoben werden, die für die zentralen Fragestellungen der Studien relevant sind. Es gibt jedoch wenig zusätzlichen Platz in den Erhebungsinstrumenten und entsprechend ist die Verwendung möglichst sparsamer Messverfahren von Vorteil.

Daher wurde basierend auf früheren Konstruktionsschritten und unter Berücksichtigung der Itemkennwerte aus vorherigen Studien eine Kurzskala PREIS-K entwickelt (Lüke und Grosche 2020). Zudem wurde die Instruktion deutlich gekürzt.

Bei der Auswahl der Items wurde besonderer Wert auf weitgehende Erhaltung der Inhaltsvalidität gelegt, auch wenn ggf. einige der beibehaltenen Items in der Gesamtskala weniger trennscharf waren als andere. Die Kurzskaala wurde nicht mit dem Ziel konstruiert, vollwertige Einstellungsskalen zu ersetzen, sondern eine Alternative anzubieten, wenn wenig Platz bzw. Befragungszeit vorhanden ist.

Da für die originale Skala die Eindimensionalität bereits nachgewiesen werden konnte, sollte in dieser Studie die Faktorenstruktur auch für die Kurzform bestätigt und gegen ein Bifaktor-Modell getestet werden. Im einfaktoriellen Modell laden alle sechs Items auf eine latente Variable „Einstellung“. Da die Richtung der Itemformulierung (drei positiv formulierte und drei negativ formulierte Items) sich möglicherweise als Methodenfaktor auswirkt, beinhaltet das Bifaktor-Modell einen latenten inhaltlichen Faktor und zwei latente methodische Faktoren (positive & negative Itemformulierung), wobei die latenten Variablen unkorreliert sind.

Zur Prüfung der Konstruktvalidität wurden die folgenden Variablen herangezogen, da sie im Rahmen der Theorie des geplanten Verhaltens (vgl. z. B. Ajzen 2011) und durch eine aktuelle Meta-Analyse (Yada et al. 2022) empirisch belegte positive Zusammenhänge mit Einstellungen erwarten lassen: spezifische Selbstwirksamkeitserwartung hinsichtlich der Gestaltung inklusiven Unterrichts, Gefühl des Vorbereitenseins auf den gemeinsamen Unterricht, Kooperation der Lehrkräfte und Einsatz verschiedener Formen von Co-Teaching, sowie die Einschätzung zur Umsetzung von Differenzierung und Barrierefreiheit.

Für die Selbstwirksamkeit kam eine modifizierte Fassung der Subskala „Selbstwirksamkeit bezogen auf die Gestaltung inklusiven Unterrichts“ aus dem KIESEL (Bosse und Spörer 2014) zum Einsatz, mit einer internen Konsistenz von Cronbach's $\alpha=0,84$. Zudem wurde anhand eines einzelnen Items, entnommen aus dem Bildungstrend 2016 (Schipolowski et al. 2019), die Einschätzung der Lehrkräfte erfragt, wie gut sie sich insgesamt vorbereitet fühlen, im gemeinsamen Lernen zu unterrichten.

Zur Kooperation wurde zunächst anhand eines einzelnen Items erfragt, ob und mit wem eine Lehrkraft im Unterricht kooperierte. Weiter wurde ein in Anlehnung an Friend et al. (2010) entwickeltes Instrument zum Co-Teaching (Schledjewski et al. 2022) eingesetzt, das erfragte, wie oft (,nie', ,selten', ,mehrmals pro Monat', ,mehrmals pro Woche') die insgesamt sechs bezeichneten Formen des Co-Teachings im Unterricht umgesetzt werden, darunter z. B. ,teach-and-assist', ,Parallelunterricht' oder ,Teamteaching'. Es sind jedoch vermutlich nicht alle Co-Teaching Formen gleichermaßen zuträglich für gelungenen inklusiven Unterricht. So ist aus inklusionstheoretischer Sicht z. B. ,Teamteaching' besonders wünschenswert, wohingegen eine reine Beobachtung des allgemeinen Unterrichts durch eine Lehrkraft (,teach-and-observe') eher weniger geeignet sein dürfte (Paulsrud und Nilholm 2020). Im Sinne einer diskriminanten Validität sind hier eher niedrigere Zusammenhänge zu erwarten als bei inklusionsförderlichen Co-Teaching Formen.

Die Umsetzung von inklusionsförderlichen Unterrichtspraktiken wurde in Anlehnung an Theisel (2015) mit einer Skala von insgesamt zehn Items zur inneren Differenzierung im Unterricht erhoben, z. B. ,Wenn ich ein Arbeitsblatt vorbereite, entwerfe ich zwei oder mehr Versionen mit unterschiedlichem Niveau'. Die interne Konsistenz dieser Skala ist mit Cronbach's $\alpha=0,70$ für den Einsatz im Rahmen

dieser Untersuchung ausreichend. Schließlich wurde die Umsetzung von Barrierefreiheit im Unterricht in Anlehnung an eine von Lüke (2019) entwickelte Kurzskala ($k=10$ Items) gemessen (vgl. Capp 2017). Ein Beispiel-Item lautet „Ich biete dieselbe Information in meinem Unterricht auf unterschiedliche Art und Weise an (z. B. auditiv, visuell)“. Diese Skala wurde bislang noch nicht empirisch evaluiert, eine interne Konsistenz von Cronbach's $\alpha=0,75$ rechtfertigt aber ihren Einsatz als Variable zur Validierung.

1.2 Fragestellungen

Anhand der Eltern- und Lehrkraftbefragungen der Studie zur Inklusion in der Sekundarstufe I in Deutschland (INSIDE) untersuchen wir, ob die empirisch bewährte Faktorenstruktur und die positiv bewertete Messgüte der vollständigen PREIS-Skala bei der Kurzskala (PREIS-K) erhalten bleibt. Es werden die folgenden drei Teilfragenstellungen bearbeitet:

1. Ist die Messstruktur eindimensional, wie angenommen, oder passt ein Bifaktor-Modell, das die Polung der Itemformulierung berücksichtigt, besser zu den empirischen Daten? Hier erwarten wir, dass aufgrund der Aufteilung in positiv und negativ formulierte Items ein Bifaktor-Modell mit einem Methodenfaktor eine signifikant höhere Modellpassung aufweist.
2. Ist die Messstruktur in den beiden Gruppen der Lehrkräfte und Eltern invariant? Ist die Messstruktur in den beiden Gruppen der Regelschullehrkräfte und der sonderpädagogischen Lehrkräfte invariant? Hier erwarten wir angesichts der auf das Schulsystem bezogenen Items, dass sich sowohl für den Vergleich der Lehrkräfte mit den Eltern als auch für den Vergleich innerhalb der Lehrkräfte eine invariante Messstruktur zeigt.
3. Finden sich in den Korrelationen mit anderen Konstrukten (auf Ebene der Lehrkräfte) Hinweise gegen die konvergente Validität? Hier erwarten wir positive Korrelationen mit konvergenten Variablen wie der Selbstwirksamkeit oder dem dem gezeigten Lehrkraft-Verhalten.

2 Methode

Das Design und die geplanten Analysen dieser Studie wurden vor der Bereitstellung der Forschungsdaten präregistriert (<https://doi.org/10.17605/OSF.IO/D546G>). Etwaige Anpassungen und Abweichungen vom geplanten Vorgehen werden explizit benannt.

2.1 Stichprobe

Datengrundlage bildet die erste Kohorte einer bundesweiten Längsschnittstudie des Projekts INSIDE, die im Frühjahr 2019 an allgemeinen Schulen in Deutschland erhoben wurde. Es wurden Klassenlehrkräfte, Fachkräfte, sonderpädagogische Lehrkräf-

te, Schulleitungen und Eltern befragt. Der gesamte Datensatz wird nach Abschluss des Projektes öffentlich verfügbar gemacht.

Es liegen Antworten von insgesamt $N=1038$ Lehrkräften und $N=1438$ Eltern vor. Alle Personen, die weniger als fünf der sechs Items der PREIS-K beantwortet hatten (Lehrkräfte: $n=31$ und Eltern: $n=11$) oder im Rahmen einer Plausibilitätsprüfung auffällige Antwortmuster aufwiesen (identische Antworten trotz positiv und negativ gepolter Items; Lehrkräfte: $n=5$ und Eltern: $n=4$), wurden aus den Analysen ausgeschlossen. Es verblieben somit die Antworten von $n=1002$ Lehrkräften und $n=1423$ Eltern im Datensatz. Von den Lehrkräften unterrichteten $n=797$ als allgemeine und $n=205$ als sonderpädagogische Lehrkraft. Die Lehrkräfte waren zu 76 % weiblich, wobei der Anteil an Frauen bei den sonderpädagogischen Lehrkräften (83,8 %) im Vergleich zu den Regelschullehrkräften (75 %) leicht überwog ($p=0,023$); bei den befragten Eltern gaben 82 % als Geschlecht weiblich an. Die Lehrkräfte waren im Durchschnitt 42,5 Jahre alt ($SD=10,9$) und konnten auf $M=13$ Jahre Unterrichtserfahrung zurückblicken ($SD=10,9$), wobei sich Regelschul- und sonderpädagogische Lehrkräfte nicht signifikant voneinander unterschieden (Alter $p=0,34$ bzw. Berufsjahre $p=0,329$).

2.2 Material und Design

Die gekürzte Fassung (PREIS-K) unterscheidet sich von der ursprünglichen PREIS-Skala in den folgenden drei Aspekten:

- Minimale Veränderungen an Itemformulierungen, die uns im Zuge der ministeriellen Genehmigungsverfahren zurückgemeldet wurden (Änderungen siehe Tab. 7 im Anhang).
- Deutliche Kürzung und leichte inhaltliche Anpassungen der Instruktion, um die Testökonomie zu sichern.
- Reduktion des Antwortformats von einer fünf- auf eine vierstufige Likert-Skala, zur Anpassung an die in der INSIDE-Studie verwendeten vierstufigen Skalen („stimme überhaupt nicht zu“, „stimme eher nicht zu“, „stimme eher zu“, „stimme voll und ganz zu“).

In Tab. 7 im Anhang werden Instruktionstext und Wortlaut der Items aus der originalen und der verkürzten Version gegenübergestellt.

Bei der Befragung zur Kooperation wurde das dreistufige Antwortformat („nein“, „mit Regelschullehrkraft“, „mit Sonderpädagog*in“) zu einer dichotomen Variable zusammengefasst („Kooperation findet statt: ja/nein“). Bei den drei Variablen Kooperation, Barrierefreiheit und Differenzierung lagen für einzelne Lehrkräfte teilweise mehrere Angaben vor, die auf Personenebene zusammengefasst wurden. Hintergrund ist, dass einzelne Lehrkräfte in mehreren Erhebungsklassen und verschiedenen Fächern bzw. Rollen (in Deutsch, Mathematik und als Klassenlehrkräfte) arbeiten und die Fragen pro Klasse ausfüllten. Da bei der vorliegenden Fragestellung die Ausprägung auf Ebene der Person der Lehrkräfte im Vordergrund stand, erschien es uns gerechtfertigt, diese Werte zu einer durchschnittlichen Ausprägung der einzelnen Lehrkräfte zusammenzufassen.

2.3 Auswertungen

Das Problem von einzelnen fehlenden Werten (bei 29 Personen fehlte ein einzelnes Item) wurde über eine FIML-Schätzung adressiert; in einigen Fällen sorgte diese Prozedur jedoch dafür, dass die Modelle nicht konvergierten, was sich z.T. durch einen listenweisen Fallausschluss beheben ließ.

Die Modelle wurden mit R (R Core Team 2022) unter Verwendung von lavaan (Rosseel 2012) geschätzt. Es wurde der robuste Schätzer MLR verwendet; die Modellvergleiche erfolgten anhand der klassischen ML-Schätzer. Zu jedem Modell werden die folgenden Indices berichtet: χ^2 (df), CFI, TLI, RMSEA, SRMR, AIC und BIC. Als Kriterium für die Modellpassung werden die folgenden Grenzwerte angelegt, orientiert an den gängigen Konventionen (Backhaus et al. 2013): Der χ^2 -Test ist nicht signifikant (d. h. die aus den Modellparametern resultierende theoretische Varianz-Kovarianzmatrix weicht nicht signifikant von der empirischen Varianz-Kovarianzmatrix ab), CFI und TLI sollten Werte größer als 0,95 aufweisen, RMSEA sollte kleiner als 0,08 und SRMR kleiner als 0,06 sein. Dabei orientieren wir uns bei der Entscheidung für die Annahme des Modells nicht an einem einzelnen Wert, sondern am Gesamtbild der Indizes (z. B. ist der χ^2 -Test bei großen Stichproben sehr sensitiv).

Das in Fragestellung 1 als passend identifizierte Modell wurde gemäß den gängigen vier Stufen der Messinvarianz überprüft (Putnick und Bornstein 2016): konfigurale Messinvarianz mit gleicher Faktorenstruktur, metrische Messinvarianz mit gleicher Ladungsstruktur, skalare Messinvarianz mit gleicher Struktur der Intercepts und residuale Messinvarianz mit gleicher Struktur der Messfehler. Die Messinvarianz wurde zunächst zwischen den Gruppen der Lehrkräfte und Eltern geprüft. Anschließend wurden auf Lehrkraft-Ebene die Regelschullehrkräfte mit den sonderpädagogischen Lehrkräften verglichen. Als zusätzlicher explorativer Vergleich (aufgrund deutlich geringerer Teilstichproben) wurden Lehrkräfte mit und ohne Lehramtsstudium (Quereinsteiger*innen) verglichen, um die Messäquivalenz über unterschiedliche Professionalisierungsstufen hinweg zu überprüfen. Falls die Modellpassungen nicht ausreichend gut waren, wurden nachträglich auch Modifikations-Indices betrachtet. Falls es inhaltlich plausibel und vertretbar war, wurden die entsprechenden Parameter anschließend frei geschätzt. Sämtliche Entscheidungen legen wir im Ergebnisteil transparent dar.

Schließlich wurde auf Lehrkraft-Ebene die Konstruktvalidität durch Zusammenhänge mit Selbstwirksamkeit, Gefühl des Vorbereitet-Seins auf Gemeinsamen Unterricht, Unterrichten im Team, Umsetzung von Barrierefreiheit und Differenzierung überprüft. Konvergente Validität nehmen wir an, wenn die Einstellungen weder zu hoch ($r > 0,50$) noch zu niedrig ($r < 0,30$) mit diesen Variablen korrelieren.

3 Ergebnisse

3.1 Fragestellung 1 – Modellüberprüfung

Die deskriptiven Itemstatistiken pro Subgruppe (Lehrkräfte und Eltern) sind in Tab. 1 angeführt, sie umfassen die Itemschwierigkeit (M) und Standardabweichung (SD), sowie Skewness, Kurtosis und die Trennschärfen der Items. Jedes Item wurde von mindestens 99 % der Probanden beantwortet; aufgrund der Datenbereinigung wurden nur Probanden berücksichtigt, die mindestens fünf Items beantwortet hatten.

Die interne Konsistenz der Kurzskaala ist mit Cronbachs $\alpha=0,92$, 95 % KI [0,91, 0,93] bei den Lehrkräften und mit $\alpha=0,90$ [0,89, 0,91] bei den Eltern hoch. Auch innerhalb der Lehrkraftgruppen ist eine hohe interne Konsistenz gegeben: $\alpha=0,91$ [0,91, 0,92] für Regelschullehrkräfte und $\alpha=0,90$ [0,88, 0,92] für sonderpädagogische Lehrkräfte. Die Korrelationen der Items untereinander sind in Tab. 2 abgetragen.

Die Passung des eindimensionalen Modells sowohl übergreifend für die gesamte Stichprobe als auch für die Lehrkräfte und Eltern getrennt ist mit Ausnahme des RMSEA zufriedenstellend (siehe Tab. 3). Jedoch führte das Bifaktor-Modell zu einer signifikant besseren Passung: Für die gesamte Stichprobe $\Delta\chi^2(6)=265,01$, $p<0,001$; sowie getrennt für Lehrkräfte $\Delta\chi^2(6)=112,38$, $p<0,001$; und Eltern $\Delta\chi^2(6)=144,27$, $p<0,001$. Bei der gesamten Stichprobe lagen die Faktorladungen für den generellen (inhaltlichen) Faktor dabei mit Werten zwischen $\beta=0,733$ und $0,831$ deutlich höher als die Faktorladungen für die methodischen Faktoren (positiv/negativ) mit Werten zwischen $\beta=0,126$ und $0,356$; bei den Subgruppen der Lehrkräfte und Eltern zeigte sich ein ähnliches Muster. Auch eine Betrachtung des

Tab. 1 Deskriptive Itemstatistiken der PREIS-K Items, getrennt nach Gruppen (Lehrkräfte und Eltern)

Items	Lehrkräfte gesamt				
	M	SD	Skewness	Kurtosis	Trennschärfe
PR01	2,52	0,91	0,03	-0,81	0,78
PR02r	2,48	1,06	0,07	-1,22	0,82
PR03r	2,33	0,95	0,20	-0,89	0,77
PR04r	2,76	0,97	-0,24	-0,98	0,75
PR05	2,31	0,92	0,25	-0,75	0,77
PR06	2,34	0,97	0,21	-0,93	0,72
Items	Eltern				
	M	SD	Skewness	Kurtosis	Trennschärfe
PR01	2,88	0,87	-0,38	-0,56	0,74
PR02r	2,91	1,10	-0,47	-1,17	0,66
PR03r	2,57	1,08	-0,11	-1,25	0,73
PR04r	3,09	1,00	-0,74	-0,63	0,71
PR05	2,91	0,98	-0,45	-0,88	0,78
PR06	2,89	0,98	-0,42	-0,92	0,76

Vierstufiges Antwortformat, Wertebereich 1 bis 4. Items PR02r, PR03r und PR04r sind negativ formuliert und wurden uncodiert. Die Items im Wortlaut finden sich in Tab. 7 im Anhang sowie bei Lücke und Grosche (2020)

Tab. 2 Interkorrelationen der PREIS-K Items, getrennt nach Lehrkräften und Eltern

	PR01	PR02r	PR03r	PR04r	PR05	PR06
PR01	–	0,714	0,647	0,630	0,674	0,637
PR02r	0,528	–	0,717	0,699	0,671	0,653
PR03r	0,579	0,595	–	0,713	0,634	0,559
PR04r	0,555	0,568	0,624	–	0,594	0,551
PR05	0,707	0,540	0,602	0,606	–	0,698
PR06	0,698	0,524	0,590	0,588	0,749	–

Items PR02r, PR03r und PR04r sind negativ formuliert und wurden umcodiert. Alle Korrelationen sind mit $p < 0,001$ signifikant. Werte oberhalb der Diagonalen beziehen sich auf die Stichprobe der Lehrkräfte, unterhalb der Diagonalen auf Eltern

Tab. 3 Modellvergleich eindimensionales vs. Bifaktor-Modell über alle Personen, sowie nach Personengruppen der Lehrkräfte und Eltern getrennt

Modell	χ^2 (df)	p (χ^2)	CFI	TLI	RMSEA	SRMR	AIC	BIC
Gesamt ein-dimensional	253,042 (9)	<0,001	0,963	0,938	0,125 [0,112; 0,138]	0,030	32.549	32.653
Gesamt Bifaktor	2,839 (3)	0,417	1,000	1,000	0,000 [0,000; 0,042]	0,003	32.212	32.351
Lehrkräfte ein-dimensional	106,557 (9)	<0,001	0,968	0,947	0,120 [0,100; 0,141]	0,032	12.318	12.377
Lehrkräfte Bifaktor	5,552 (3)	0,136	0,999	0,995	0,037 [0,000; 0,085]	0,006	12.201	12.289
Eltern ein-dimensional	139,882 (9)	<0,001	0,962	0,937	0,120 [0,103; 0,138]	0,034	19.382	19.477
Eltern Bifaktor	2,917 (3)	0,405	1,000	1,000	0,000 [0,000; 0,055]	0,004	19.200	19.327

Robuste Schätzer (MLR). „Gesamt“ ist der vollständige bereinigte Datensatz ($N = 2425$), bestehend aus Lehrkräften ($n = 1002$) und Eltern ($n = 1423$)

Anteils der erklärten gemeinsamen Varianz (Rodriguez et al. 2016) zeigt, dass 88 % der Varianz auf den generellen Faktor zurückzuführen sind.

3.2 Fragestellung 2 – Messinvarianz zwischen Subgruppen

Die Ergebnisse für den Vergleich der Bifaktor-Modelle von Eltern und Lehrkräften sind in Tab. 4, der Vergleich innerhalb der Lehrkräfte (Regelschullehrkräfte und sonderpädagogische Lehrkräfte) in Tab. 5 zu sehen. Bei der Messinvarianz zwischen Lehrkräften und Eltern zeigt einerseits die CFI-Differenz, dass für die metrische ($\Delta CFI = 0,009$) und skalare Stufe ($\Delta CFI = 0,004$) das Kriterium von $\Delta CFI < 0,01$ (Putnick und Bornstein 2016) erfüllt ist. Insofern kann davon ausgegangen werden, dass über die Faktorenstruktur auch die Ladungsmuster und Intercepts zwischen

Tab. 4 Messinvarianz für das Bifaktor-Modell im Vergleich zwischen Lehrkräften und Eltern

LK vs. Eltern (Bifaktor)	df	AIC	BIC	χ^2	$\Delta\chi^2$	Δdf	p-Wert	CFI	ΔCFI
Konfigural	12	31.774	32.017	49,471	–	–	–	0,996	–
Metrisch	19	31.841	32.043	130,496	69,745	7	<0,001	0,987	0,009
Skalar	23	31.875	32.055	172,798	40,381	4	<0,001	0,983	0,004
Residual	29	32.016	32.160	325,433	101,959	6	<0,001	0,967	0,016

Tab. 5 Messinvarianz für das Bifaktor-Modell im Vergleich zwischen Regelschullehrkräften und Sonderpädagog:innen

Regelschule vs. Sonderpäd. (Bifaktor)	df	AIC	BIC	χ^2	$\Delta\chi^2$	Δdf	p-Wert	CFI	ΔCFI
Konfigural	6	12.491	12.727	15,405	–	–	–	0,998	–
Metrisch	15	12.478	12.669	19,923	4,5179	9	0,874	0,999	–0,001
Skalar	18	12.472	12.649	20,716	0,7934	3	0,851	0,999	0,000
Residual	24	12.471	12.618	31,553	10,836	6	0,094	0,995	0,004

den Gruppen vergleichbar sind. Andererseits spricht ein signifikanter $\Delta\chi^2$ -Wert dafür, dass die Annahme der Messäquivalenz über das konfigurale Level hinaus abgelehnt werden muss, da bereits die Aufnahme der metrischen Restriktion zu einer signifikanten Verschlechterung der Modellpassung führt. Bei dieser konservativen und stichprobengrößenabhängigen Interpretation basierend auf dem χ^2 -Test ist somit lediglich die Faktorenstruktur zwischen den Gruppen vergleichbar, die Ladungsmuster, Intercepts und Residuen dagegen nicht. Die nachfolgende Überprüfung anhand der Modifikationsindizes ergab, dass auch durch mögliche Lockerungen der Restriktion für bis zu drei Ladungsparameter keine partielle Messinvarianz erreicht werden konnte.

Innerhalb der Lehrkräfte ist die Skala für Regelschul- und sonderpädagogische Lehrkräfte auf metrischer, skalarer und residualer Ebene ($p=0,874; 0,851; 0,094$) messinvariant. Somit sind über die grundlegende faktorielle Struktur hinaus auch die Ladungsmuster, die Intercepts und die Messfehler zwischen den Gruppen vergleichbar. Jedoch resultierte die Aufnahme der skalaren Restriktion in einem Warnhinweis, dass das Modell nicht fehlerfrei konvergiert sei und nur auf Kosten eines negativ geschätzten Parameters konvergierte; dieses Ergebnis muss also mit der gebotenen Zurückhaltung interpretiert werden. Konservativ interpretiert ist aber auf jeden Fall eine metrische Messinvarianz erfüllt.

Bei dem explorativen Vergleich von Lehrkräften mit Lehramtsstudium und Quereinsteiger*innen ließen sich die Bifaktor-Modelle aufgrund von Konvergenzproblemen nicht schätzen. Daher wird an dieser Stelle abweichend von der Präregistrierung die Messinvarianz für das einfaktorielle Modell berichtet. Für diesen explorativen Vergleich bei einem eindimensionalen Modell kann die Messinvarianz auf metrischer, skalarer und residualer Ebene ($p=0,635; 0,761; 0,289$) angenommen werden. Die vollständige Tab. 10 zur Analyse der Messinvarianz für Lehrkräfte, die Lehramt studiert haben, im Vergleich mit Quereinsteiger*innen ist im Anhang zu finden.

Tab. 6 Übersicht Variablen konvergente Validität

Konstrukt	Itemanzahl und Quelle	r	p	Koeffizient
<i>Selbstwirksamkeit</i>	$k=7$, Bosse und Spörer (2014)	0,45	<0,001	Pearson
<i>Vorbereitet-sein GU</i>	$k=1$, Schipolowski et al. (2019)	0,38	<0,001	Punkt-biserial
<i>Kooperation allgemein</i>	$k=1$	0,18	<0,001	Punkt-biserial
Teach-observe	$k=6$ Einzelitems, Schledjewski et al. (2022)	-0,01	0,90	Spearman
Teach-assist		0,00	0,97	Spearman
Niveaudifferenzierung		0,13	0,01	Spearman
Parallelunterricht		0,05	0,39	Spearman
Stationsunterricht		0,09	0,10	Spearman
Teamteaching		0,24	<0,001	Spearman
<i>Umsetzung UDL</i>	$k=10$, Lücke (2019)	0,15	<0,001	Pearson
<i>Differenzierung</i>	$k=10$, angel. an Theisel (2015)	0,26	<0,001	Pearson

3.3 Fragestellung 3 – externe Validität

Die PREIS-K korreliert mit der Selbstwirksamkeit in Höhe von $r=0,45$, $p<0,001$. Für den ordinalskalierten Indikator „Vorbereitet sein auf den GU“ zeigte sich ein Zusammenhang von $r_{\text{Spearman}}=0,38$, $p<0,001$. Die beiden Drittvariablen „Selbstwirksamkeit“ und „Vorbereitet sein“ korrelierten untereinander mit einer mittleren Effektstärke $r_{\text{Spearman}}=0,47$, $p<0,001$. Bei der allgemeinen Erfassung, ob Kooperation im Unterricht umgesetzt wurde, ergab sich eine signifikante punkt-biserial Korrelation von $r=0,18$, $p<0,001$. Die PREIS-K korrelierte schwach positiv mit den inklusionsrelevanten Formen des Co-Teachings ‚Niveaudifferenzierung‘ ($r_{\text{Spearman}}=0,13$, $p=0,01$) und ‚Teamteaching‘ ($r_{\text{Spearman}}=0,24$, $p<0,001$), jedoch ohne das vorher festgelegte Kriterium von $r>0,30$ zu erreichen. Bei Co-Teaching-Formen, die inklusionstheoretisch kritisiert werden, ließen sich hypothesenkonform keine signifikanten Zusammenhänge feststellen. Die Umsetzung von Barrierefreiheit korrelierte mit $r=0,15$, $p<0,001$; die Berücksichtigung von Heterogenität durch differenzierende Maßnahmen mit $r=0,26$, $p<0,001$ (Tab. 6).

4 Diskussion

In der vorliegenden Studie wurden die psychometrischen Eigenschaften der aus sechs Items bestehenden Kurzfassung der PREIS-Skala an einem großen Datensatz der INSIDE-Studie überprüft, um sicherzustellen, dass die Testgüte auch bei einer verkürzten Version erhalten bleibt.

Die deskriptiven Ergebnisse zu einzelnen Items wiesen auf symmetrische Verteilungen hin, die Mittelwerte lagen nahe bei der theoretischen Skalenmitte von 2,5 und die Standardabweichung betrug bei allen Items etwa einen Skalenpunkt. Lediglich bei den Eltern gab es eine leichte Zustimmungstendenz bei einem Item. Die Interkorrelationen zwischen den Items lagen durchschnittlich bei $r=0,65$ (Lehrkräfte) bzw. $r=0,60$ (Eltern) und allesamt in einem zufriedenstellenden Bereich von mindestens $r>0,50$. Die Trennschärfen lagen mit Werten zwischen $r=0,72$ bis 0,82

(Lehrkräfte) bzw. $r=0,66$ bis $0,78$ (Eltern) in einem angemessenen Bereich. Zudem kann die interne Konsistenz der PREIS-K mit Werten von $\alpha=0,92$ (Lehrkräfte) bzw. $\alpha=0,90$ (Eltern) als hoch eingestuft werden.

Bei der ersten Fragestellung wurde geprüft, ob ein eindimensionales Messmodell oder ein Bifaktor-Modell mit Methodenfaktor besser zu den Daten passt. Das einfaktorielle Modell wies bereits eine zufriedenstellende Modellpassung auf (lediglich der RMSEA verfehlte den vorab definierten Zielbereich). Das signifikant besser passende Bifaktor-Modell deutet darauf hin, dass die unterschiedliche Polung der Itemformulierungen bei der Auswertung berücksichtigt werden könnte. Jedoch sind die Faktorladungen des Generalfaktors durchschnittlich deutlich höher als die Ladungen der Methodenfaktoren. Auch beträgt der Anteil der durch den generellen Faktor aufgeklärten Varianz sowohl für das Gesamtmodell mit Eltern und Lehrkräften als auch für das Modell ausschließlich mit Lehrkräften 88% , weshalb auch die sparsamere einfaktorielle Interpretation möglich ist (Rodriguez et al. 2016) und aus inhaltlich-theoretischen Gründen auch deutlich sinnvoller erscheint. Der zusätzliche Mehrwert der Berücksichtigung der Methodenfaktoren dürfte für die Praxis damit überschaubar bleiben. Demnach bewerten wir die PREIS-K als eindimensional, was zu der bisherigen Studienlage für die Langversion (PREIS) passt. Für die meisten Forschungsdesigns dürfte die einfaktorielle Auswertung ohnehin die beliebtere Variante sein, da die Auswertung und Interpretation der gewonnenen Einstellungsdaten klarer und weniger aufwändig ist.

Die zweite Fragestellung lautete, ob die PREIS-K zwischen verschiedenen Gruppen identisch zu interpretieren ist oder ob Gruppenvergleiche unzulässig sind. Beim Vergleich von Lehrkräften und Eltern war bei Berücksichtigung der relativen Passung (ΔCFI) die skalare Stufe der Messinvarianz erfüllt. Bei ausschließlicher Berücksichtigung des konservativen und stichprobenabhängigen χ^2 -Tests wären nur die Faktorenstruktur beider Gruppen, nicht aber die Ladungsmuster etc. vergleichbar. Aufgrund ihrer unterschiedlichen Rollen im Erziehungskontext und der eigenen professionellen Involviertheit als Lehrkraft ist es plausibel, dass Eltern sich von Lehrkräften systematisch darin unterscheiden, wie sie Items aus der PREIS-K interpretieren. Angesichts der hohen CFI-Werte und der großen Stichprobe, die eine möglicherweise übermäßige Sensitivität des χ^2 -Tests bewirkt, bezeichnen wir direkte Vergleiche der Einstellungen beider Gruppen als durchaus möglich.

Beim Vergleich innerhalb der Lehrkräfte zwischen Regelschul- und sonderpädagogischen Lehrkräften kann dagegen vollständige (residuale) Messinvarianz angenommen werden, wenn auch aufgrund der hohen Voraussetzungen der Bifaktor-Modelle ein Warnhinweis in der Analyse auftrat. Eine nachträgliche explorative Analyse des weniger komplexen einfaktoriellen Modells bestätigte die residuale Messinvarianz (vgl. Tab. 9 im Anhang). Der explorative Vergleich zwischen Lehrkräften mit Lehramtsstudium und Quereinsteiger*innen erfüllte ebenfalls residuale Messinvarianz. Somit werden die Items der PREIS-K zur Erfassung von Einstellungen zu einem inklusiven Schulsystem bei Lehrkräften unabhängig von deren Ausbildung auf eine ähnliche Weise verstanden. Vergleiche zwischen verschiedenen Gruppen von Lehrkräften sind folglich zulässig und sinnvoll interpretierbar.

Die dritte Fragestellung kann weniger eindeutig beantwortet werden. Für Variablen, die recht nah am Konstrukt der Einstellungen sind, wie z. B. die Selbstwirk-

samkeit, ließ sich ein deutlicher hypothesenkonformer Zusammenhang nachweisen: So sind die Einstellungen der Lehrkräfte zu einem inklusiven Schulsystem und ihre spezifischen Selbstwirksamkeitserwartungen in diesem Bereich mit $r=0,45$ deutlich assoziiert. Dies entspricht den zu erwartenden Zusammenhängen dieser Variablen (metaanalytisch zwischen $r=0,31$ und $0,39$; Yada et al. 2022). Der Zusammenhang mit dem Gefühl des Vorbereitet-Seins auf den Gemeinsamen Unterricht zeigt einen ähnlichen plausiblen Zusammenhang mit den Einstellungen ($r_{\text{Spearman}}=0,38$).

Dagegen zeigten sich bei den eher auf das resultierende Verhalten abzielenden Variablen, wie z. B. die konkrete Umsetzung von Barrierefreiheit nur geringere Zusammenhänge, die das vorab festgelegte Kriterium von $r>0,30$ nicht erfüllten. Bei der Umsetzung von Differenzierungsmaßnahmen ($r=0,26$) und der Co-Teaching Form des Team-Teachings ($r=0,24$) gilt dies gleichermaßen. Tendenziell weisen diese Ergebnisse aber in die richtige Richtung, da diese Formen inklusionstheoretisch besonders wünschenswert sind (Paulsrud und Nilholm 2020). Andere Formen, wie z. B. reine Beobachtung des allgemeinen Unterrichts durch eine Lehrkraft, sind vermutlich weniger für die Implementation von schulischer Inklusion geeignet (ebd.). Die fehlende Korrelation mit der PREIS-K spricht daher für deren Validität.

Als inhaltliche Limitation unserer Studie ist zunächst festzustellen, dass die PREIS-K sich auf Einstellungen zu einer komplexen Definition von Inklusion bezieht, nämlich auf die Erfüllung der Bedürfnisse aller Schüler*innen (Göransson und Nilholm 2014). Sollen Einstellungen zu einem inklusiven Schulsystem im Sinne der Platzierung von Lernenden mit Behinderungen in Regelschulen oder zum Umgang mit Behinderungen im Unterricht erfasst werden, ist die PREIS-K ungeeignet. Weiterhin ist anzumerken, dass in der Theorie des geplanten Verhaltens (Ajzen 2011) die hier gemessenen allgemeinen Einstellungen nur einen von mehreren Verhaltensprädiktoren darstellen; für die Prädiktion des tatsächlichen Verhaltens können sich somit andere Faktoren situationsabhängig deutlich stärker auswirken. So können hinsichtlich der Verhaltensintention und des tatsächlichen Verhaltens beispielsweise geringe Selbstwirksamkeitserwartungen aufgrund gemachter negativer Erfahrungen eine positive Einstellung zu einem inklusiven Schulsystem überlagern.

Eine methodische Limitation ergibt sich aus den teilweise vorhandenen Konvergenzproblemen der Bifaktor-Modelle. Zwar sind die Modelle inhaltlich sehr gut interpretierbar, sie sollten jedoch an unabhängigen Datensätzen repliziert werden. Außerdem sind systematische Verzerrungen durch die Stichprobenziehung bzw. eine geringe Rücklaufquote nicht auszuschließen. So beteiligte sich nur ein geringer Anteil der kontaktierten Schulen, weshalb ergänzend weitere Schulen aus der Grundgesamtheit angeschrieben wurden.

Insgesamt hat sich die PREIS-K im hier dargestellten Kontext einer schriftlichen Befragung von Lehrkräften und Eltern im Rahmen einer groß angelegten Querschnittsstudie bewährt. Wenn Einstellungen zu Inklusion nicht zu den Hauptfragestellungen eines Forschungsprojekts gehören, können wir die Nutzung der PREIS-K empfehlen. Die erhebliche Reduktion der Itemzahl hat aber selbstverständlich ihren Preis: Auch wenn die Auswahl der Items theoriegeleitet erfolgte, so wird eben doch ein deutlich kleinerer Ausschnitt des Konstrukts *Einstellungen zu schulischer Inklusion* abgedeckt. Falls also die Einstellungen von zentraler Bedeutung in einem

Forschungsprojekt sind, halten wir den Einsatz der vollständigen PREIS oder eines anderen empirisch bewährten Instruments für angemessener.

Die Items der PREIS-K haben den üblichen kultusministeriellen Autorisierungsprozess der Bundesländer bereits erfolgreich durchlaufen; hierauf kann im Genehmigungsverfahren von zukünftigen Studien verwiesen werden. Insbesondere für den Einsatz in breit angelegten Vergleichs- und Längsschnittstudien (wie z. B. Large-Scale Assessments), die zumeist eine Vielzahl von Variablen umfassen und nur wenig Testzeit bzw. Platz im Fragebogen haben, halten wir die ökonomische PREIS-K für grundsätzlich gut geeignet.

5 Anhang

Tab. 7 Gegenüberstellung der Item-Formulierungen in der originalen PREIS-Skala und der gekürzten PREIS-K

Instruktionstext	PREIS		PREIS-K	
	Nr.	Item	Nr.	Item
Es gibt viele unterschiedliche Meinungen und Definitionen dazu, was genau unter einem inklusiven Schulsystem zu verstehen ist. Eine dieser Definitionen ist die folgende: „In einem inklusiven Schulsystem besuchen alle Kinder eine gemeinsame Schule. Die Schulen bieten einen Unterricht an, der allen Kindern Lernmöglichkeiten bietet – unabhängig von ihren Lernvoraussetzungen, Erfahrungen und Bedürfnissen.“ Bitte denken Sie bei der Beantwortung der Fragen nur an so eine Art eines inklusiven Schulsystems, auch wenn Sie persönlich das anders sehen. Wie sehr stimmen Sie den folgenden Aussagen zu?	01	Ich vermute, dass ein inklusives Schulsystem den Unterricht für alle Kinder verbessern könnte	1	Ich vermute, dass ein <i>solches</i> inklusives Schulsystem den Unterricht für alle Kinder verbessern könnte
	04	Bei einem Volksentscheid würde ich wahrscheinlich gegen ein inklusives Schulsystem stimmen	2	Bei einem Volksentscheid würde ich wahrscheinlich gegen ein <i>solches</i> inklusives Schulsystem stimmen
	07	Ich finde es unrealistisch, Kinder mit sehr unterschiedlichen Schwierigkeiten und Bedürfnissen in einer Klasse unterrichten zu wollen	3	Ich finde es unrealistisch, Kinder mit sehr unterschiedlichen Schwierigkeiten und Bedürfnissen in einer <i>solchen Klasse zu unterrichten</i>
	11	Ich fühle mich unwohl, wenn ich mir ein inklusives Schulsystem vorstelle	4	Ich fühle mich unwohl, wenn ich mir ein <i>solches</i> inklusives Schulsystem vorstelle
	12	Ich würde viel dafür tun, dass mein eigenes Kind in einem inklusiven Schulsystem unterrichtet wird	5	Ich würde viel dafür tun, dass mein eigenes Kind in einem <i>solchen</i> inklusiven Schulsystem unterrichtet wird
	14	Es wäre schön, wenn zukünftig alle Kinder eine gemeinsame Schule in einem inklusiven Schulsystem besuchen würden	6	Es wäre schön, wenn zukünftig alle Kinder eine gemeinsame Schule in einem <i>solchen</i> inklusiven Schulsystem besuchen würden

Tab. 8 Messinvarianz für das einfaktorielles Modell im Vergleich zwischen Lehrkräften und Eltern

Lehrkräfte vs. Eltern (einfaktoriell)	df	AIC	BIC	χ^2	$\Delta\chi^2$	Δdf	p-Wert
Konfigural	18	32.058	32.267	345,87	–	–	–
Metrisch	23	32.122	32.302	420,00	74,135	5	<0,001
Skalar	28	32.233	32.383	540,44	120,440	5	<0,001
Residual	34	32.390	32.506	710,00	169,559	6	<0,001

Tab. 9 Messinvarianz für das einfaktorielles Modell im Vergleich zwischen Regelschullehrkräften und Sonderpädagog*innen

Regelschule vs. Sonder- päd. (einfaktoriell)	df	AIC	BIC	χ^2	$\Delta\chi^2$	Δdf	p-Wert
Konfigural	18	12.609	12.785	157,00	–	–	–
Metrisch	23	12.601	12.754	159,71	2,731	5	0,741
Skalar	28	12.599	12.727	167,89	8,176	5	0,147
Residual	34	12.611	12.709	191,19	12,260	6	0,056

Tab. 10 Messinvarianz für das einfaktorielles Modell im Vergleich zwischen Lehrkräften mit Lehramtsstudium und Quereinsteiger*innen

Lehramt vs. Quereinstieg (einfaktoriell)	df	AIC	BIC	χ^2	$\Delta\chi^2$	Δdf	p-Wert
Konfigural	18	12.568	12.745	161,35	–	–	–
Metrisch	23	12.562	12.714	164,98	3,424	5	0,635
Skalar	28	12.554	12.682	167,55	2,602	5	0,761
Residual	34	12.557	12.655	182,32	7,354	6	0,289

Tab. 11 Übersicht über Validierungskonstrukte

Konstrukt	Itemanzahl	Quelle	Beispiel-Item
<i>Selbstwirksamkeit</i>	<i>k</i> = 7	Bosse und Spörer (2014)	„Ich bin mir sicher, dass ich auch bei größeren Leistungsunterschieden der Schülerinnen und Schüler für alle jeweils passende Methoden und Sozialformen einsetzen kann.“
<i>Vorbereitet-sein GU</i>	<i>k</i> = 1	Schipolowski et al. (2019)	„Wie gut fühlen Sie sich aufgrund Ihrer Ausbildung, Fortbildung bzw. Berufserfahrung insgesamt darauf vorbereitet ‚Gemeinsamen Unterricht‘ zu gestalten?“
<i>Kooperation allgemein</i>	<i>k</i> = 1	–	„Unterrichten Sie in dieser Klasse im Unterricht mindestens ab und zu gemeinsam mit einer anderen Lehrkraft?“
Teach-observe Teach-assist Niveaudifferenzierung Parallelunterricht Stationsunterricht Teamteaching	<i>k</i> = 6 Einzel-items	Schledjewski et al. (2022)	„Welche Formen des Co- bzw. Teamteachings werden wie oft innerhalb des Unterrichts umgesetzt?“
<i>Umsetzung UDL</i>	<i>k</i> = 10	Lüke (2019)	„Mein Unterrichtsmaterial (z. B. Arbeitsblätter, Lernkarteien) ist sprachlich und symbolisch differenziert.“
<i>Differenzierung</i>	<i>k</i> = 10	Angelehnt an Theisel (2015)	„Wenn ich ein Arbeitsblatt vorbereite, entwerfe ich zwei oder mehr Versionen mit unterschiedlichem Niveau.“

Förderung Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) unter den Förderkennzeichen IN1503A, IN1503B, IN1503C und IN1503D gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autor:innen.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- Ajzen, I. (2011). The theory of planned behavior. In P. A. M. van Lange, A. W. Kruglanski & E. T. Higgins (Hrsg.), *SAGE social psychology program. The handbook of theories of social psychology: volume one* (S. 438–459). SAGE.
- Avramidis, E., & Norwich, B. (2002). Teachers' attitudes towards integration/inclusion: a review of the literature. *European Journal of Special Needs Education, 17*(2), 129–147. <https://doi.org/10.1080/08856250210129056>.
- Backhaus, K., Erichson, B. & Weiber, R. (2013). *Fortgeschrittene Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*. Berlin: Springer.
- de Boer, A., 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. <https://doi.org/10.1080/13603110903030089>.
- Bosse, S., & Spörer, N. (2014). Erfassung der Einstellung und der Selbstwirksamkeit von Lehramtsstudierenden zum inklusiven Unterricht. *Empirische Sonderpädagogik, 6*(4), 279–299. <https://doi.org/10.25656/01:10019>.
- Capp, M. J. (2017). The effectiveness of universal design for learning: a meta-analysis of literature between 2013 and 2016. *International Journal of Inclusive Education, 21*(8), 791–807. <https://doi.org/10.1080/13603116.2017.1325074>.
- Eagly, A. H., & Chaiken, S. (2007). The advantages of an inclusive definition of attitude. *Social Cognition, 25*(5), 582–602.
- Friend, M., Cook, L., Hurley-Chamberlain, D., & Shamberger, C. (2010). Co-teaching: an illustration of the complexity of collaboration in special education. *Journal of Educational and Psychological Consultation, 20*(1), 9–27. <https://doi.org/10.1080/10474410903535380>.
- Göransson, K., & Nilholm, C. (2014). Conceptual diversities and empirical shortcomings: a critical analysis of research on inclusive education. *European Journal of Special Needs Education, 29*(3), 265–280. <https://doi.org/10.1080/08856257.2014.933545>.
- Kielblock, S., & Woodcock, S. (2023). Who's included and Who's not? An analysis of instruments that measure teachers' attitudes towards inclusive education. *Teaching and Teacher Education, 122*, 103922. <https://doi.org/10.1016/j.tate.2022.103922>.
- Lüke, T. (2019). *Kurzskala Barrierefreiheit von Unterricht – Lehrkräfteeinschätzung (KBU-L)*. <https://doi.org/10.17605/osf.io/78tjb>.
- Lüke, T., & Grosche, M. (2018). Konstruktion und Validierung der Professionsunabhängigen Einstellungsskala zum Inklusiven Schulsystem (PREIS). *Empirische Sonderpädagogik, 10*(1), 3–20. <https://doi.org/10.25656/01:15958>.
- Lüke, T., & Grosche, M. (2020). *Kurzskala zur Messung von Einstellungen zum Inklusiven Schulsystem (PREIS-K)*. <https://doi.org/10.17605/OSF.IO/DNVZT>.
- Monsen, J. J., Ewing, D. L., & Kwoka, M. (2014). Teachers' attitudes towards inclusion, perceived adequacy of support and classroom learning environment. *Learning Environments Research, 17*(1), 113–126. <https://doi.org/10.1007/s10984-013-9144-8>.
- Paulsrud, D., & Nilholm, C. (2020). Teaching for inclusion: a review of research on the cooperation between regular teachers and special educators in the work with students in need of special support. *International Journal of Inclusive Education. https://doi.org/10.1080/13603116.2020.1846799*.
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Developmental review, 41*, 71–90. <https://doi.org/10.1016/j.dr.2016.06.004>.
- R Core Team (2022). *R: A language and environment for statistical computing*. Wien: Österreich.
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: calculating and interpreting statistical indices. *Psychological methods, 21*(2), 137–150. <https://doi.org/10.1037/met0000045>.
- Röhm, A., Schnöring, A., & Hastall, M. R. (2018). Impact of single-case pupil descriptions on student teacher attitudes towards inclusive education. *Learning Disabilities: A Contemporary Journal, 16*, 37–58.
- Rosseel, Y. (2012). lavaan : an R package for structural equation modeling. *Journal of Statistical Software. https://doi.org/10.18637/jss.v048.i02*.
- Schipolowski, S., Busse, J., Rjosk, C., Mahler, N., Becker, B., & Stanat, P. (2019). *IQB-Bildungstrend 2016: Skalenhandbuch zur Dokumentation der Erhebungsinstrumente in den Fächern Deutsch und Mathematik*. Berlin: Humboldt-Universität zu Berlin, Institut zur Qualitätsentwicklung im Bildungswesen.

- Schledjewski, J., Kluge, J., Pirsch, J., & Grosche, M. (2022). Co-Teaching aus Sicht sonderpädagogischer Lehrkräfte: Nutzungshäufigkeit von Co-Teaching-Formen und das Zugehörigkeitsgefühl der sonderpädagogischen Lehrkräfte zur Schule. *Schweizerische Zeitschrift für Heilpädagogik*, 28(5–6), 36–43.
- Schulze, S., Lüke, T., Schröter, A., Krause, K., & Kuhl, J. (2019). Replikationsstudie zur Testgüte der Professionsunabhängigen Einstellungsskala zum Inklusiven Schulsystem – ein Beitrag zur Validierung. *Unterrichtswissenschaft*, 47(2), 201–219. <https://doi.org/10.1007/s42010-018-00034-3>.
- Schwarz, N., & Bohner, G. (2001). The construction of attitudes. In A. Tesser & N. Schwarz (Hrsg.), *Blackwell handbook of social psychology: intraindividual processes* (S. 436–457). Blackwell. <https://doi.org/10.1002/9780470998519.ch20>.
- Theisel, A. K. (2015). Unterrichten Sprachheilpädagogen anders? Eine vergleichende Selbsteinschätzung von Förderschullehrkräften „Sprache“ und Regelschullehrkräften zu Qualitätsmerkmalen des Unterrichts. *Empirische Sonderpädagogik*, 7(4), 320–340.
- Wilson, C., Woolfson, L. M., Durkin, K., & Elliott, M. A. (2016). The impact of social cognitive and personality factors on teachers' reported inclusive behaviour. *The British Journal of Educational Psychology*, 86(3), 461–480. <https://doi.org/10.1111/bjep.12118>.
- Yada, A., Leskinen, M., Savolainen, H., & Schwab, S. (2022). Meta-analysis of the relationship between teachers' self-efficacy and attitudes toward inclusive education. *Teaching and Teacher Education*, 109, 103521. <https://doi.org/10.1016/j.tate.2021.1035>.