

**Empirische Sonderpädagogik**, 2018, Nr. 1, S. 3-20  
ISSN 1869-4845 (Print) · ISSN 1869-4934 (Internet)

# Konstruktion und Validierung der Professionsunabhängigen Einstellungsskala zum Inklusiven Schulsystem (PREIS)

*Timo Lüke & Michael Grosche*

*Bergische Universität Wuppertal*

## Zusammenfassung

Forschungsarbeiten zum Thema Einstellungen zu Inklusion haben Hochkonjunktur. Gleichwohl muss konstatiert werden, dass viele Instrumente zur Messung dieser Einstellungen Schwächen haben, zum Beispiel: mangelnde einstellungstheoretische Fundierungen, fehlleitende Itemformulierungen, fehlende Abgrenzung von anderen Konstrukten, zu starke Fokussierung auf Lehrkräfte, Anfälligkeit für systematische Antwortverzerrungen sowie oberflächliche Validierungsstudien. Daher entwickelten wir die Professionsunabhängige Einstellungsskala zum Inklusiven Schulsystem (PREIS), die diesen Problemen adäquat begegnet. Zur Überprüfung der Testgüte führten wir vier konsekutive Studien durch. Insgesamt können wir zeigen, dass mit der PREIS erstmalig ein Instrument vorliegt, das inhaltsvalide unter Verzicht auf Formulierungen der Zweigruppen-Theorie und ohne professionsspezifisches Fachwissen vorauszusetzen, faktoriell valide und reliabel das einfaktorielle Konstrukt Einstellungen zum inklusiven Schulsystem misst. Das Instrument ist frei und offen unter einer CC-BY-SA-Lizenz verfügbar.

Schlagwörter: Einstellungen, Inklusion, Fragebogen, Bevölkerung, Testgüte, PREIS

## Construction and validation of an attitude scale measuring attitudes towards inclusive education in the population (PREIS)

### Abstract

Research on attitudes towards inclusive education is highly topical. However, it must be noted that many measures of these attitudes have weaknesses, e.g.: lack of theoretical foundations, misleading item wording, contextual overlap with other constructs, excessive focus on teachers, susceptibility to systematic bias and unconvincing validation studies. Therefore, we developed the 'Professionsunabhängige Einstellungsskala zum Inklusiven Schulsystem (PREIS)' [Attitudes Towards Inclusive Education in the Population], which adequately addresses these problems. To evaluate the scale's psychometric quality, we conducted four consecutive studies. Overall, we show that the PREIS is the first reliable and valid German instrument to measure the one-factor construct attitudes towards the inclusive school-system, without presupposing profession-specific knowledge and in the absence of two-group theory wording. The instrument is openly available under a CC-BY-SA license to promote necessary replications.

Key Words: Attitudes, Inclusion, Questionnaire, Population, Psychometric Quality, PREIS

Die Anzahl der Publikationen (und damit auch der Forschungsprojekte) über Einstellungen zu Inklusion ist in den letzten Jahren stark angestiegen. In mehreren Fachzeitschriften wurden entsprechende Themenhefte herausgegeben (u.a. Eder, Altrichter & Paseka, 2014; Schwab & Feyerer, 2016) und auch die Zahl der freien Artikel in den *peer reviewed* Fachzeitschriften ist konstant auf hohem Niveau. Die meisten Autorinnen und Autoren erklären ihr Interesse an diesem Konstrukt damit, dass positive Einstellungen zu Inklusion eine wichtige Gelingensbedingung für schulische Inklusion seien (Avramidis & Norwich, 2002; de Boer, Pijl & Minnaert, 2011). Empirische Belege für diese Annahme liegen bisher nicht vor und können prinzipiell erst dann gewonnen werden, wenn geeignete Instrumente zur Messung von Einstellungen zu Inklusion vorliegen. Im ersten Teil des vorliegenden Beitrags diskutieren wir sechs Probleme solcher Instrumente und zeigen, warum trotz vielfältiger Forschungsaktivitäten keines der aktuell vorliegenden Instrumente in der Lage ist Einstellungen zum inklusiven Schulsystem zuverlässig und valide zu messen. Im zweiten Teil beschreiben wir die Konstruktion der Professionsunabhängigen Einstellungsskala zum inklusiven Schulsystem (PREIS) und zeigen, wie wir mit dieser Neuentwicklung versuchten die sechs vorgestellten Probleme zu lösen. Die PREIS kann nun in zukünftigen Forschungsprojekten eingesetzt werden, insbesondere zur Prüfung der Annahme, ob positive Einstellungen zu Inklusion tatsächlich eine wichtige Gelingensbedingung für Inklusion sind.

### **Probleme bei der Messung von Einstellungen zum Inklusiven Schulsystem**

Der Anlass für die Neuentwicklung einer Skala zur Erfassung von Einstellungen zum inklusiven Schulsystem waren einige zentrale Schwächen der bisher verfügbaren Skalen. Diese betreffen unter anderem (1) die einstellungstheoretische Fundierung, (2) die

Formulierung der Items, (3) die Abgrenzung von anderen Konstrukten, (4) die Adressatengerichtetheit, (5) Antworttendenzen und (6) die Validierungsstudien. Eine Zusammenstellung der uns bekannten Instrumente findet sich bei Lücke, Przbilla, Krämer und Patzelt (2017). Einschränkend sei gesagt, dass dabei immer nur einzelne der Kritikpunkte auf jedes der Instrumente zutreffen.

### **Mangelnde einstellungstheoretische Fundierung**

In den vorliegenden Instrumenten zur Messung von Einstellungen zum inklusiven Schulsystem werden zum Beispiel unter Verweis auf Rosenberg und Hovland (1960) sowie Eagly und Chaiken (1993) häufig Konstrukte mit drei Komponenten (kognitiv, affektiv, behavioral) definiert. In der Testentwicklung wird dann versucht ausschließlich *eine* spezifische Komponente (meist die kognitive) zu erfassen oder die drei Komponenten in Subskalen abzubilden. Größtenteils fehlen Erklärungen dazu, wie diese Subskalen-Werte theoretisch eingebettet sind und wie sie interpretiert werden sollen. Es wird dann angenommen, dass bestimmte Itemformulierungen in der Lage wären ganz gezielt *nur eine* dieser drei Einstellungskomponenten anzusprechen. Falls die entsprechenden Skalen faktorenanalytisch untersucht wurden, waren die drei Einstellungskomponenten in der Regel nicht nachweisbar und die Faktorenstruktur wurde durch andere – meist inhaltliche – Gruppierungsfaktoren erklärt. Entweder werden dann spontan andere Subskalen gebildet oder Mischungen aus den drei Einstellungskomponenten und inhaltlichen Konzepten gewählt. Nur wenige Autorinnen und Autoren geben offen an, dass das ursprünglich erwartete Modell der drei Einstellungskomponenten auf ihr Instrument oder ihre Daten nicht passte, wie beispielsweise de Boer, Timmerman, Pijl und Minnaert (2012) sowie Seifried und Heyl (2016) bei den Validierungen ihrer Einstellungsskalen zu Inklusion.

Der Grund für die fehlende Passung des klassischen dreifaktoriellen Modells dürfte aber eher darin liegen, dass Einstellungen bzw. die Antworten auf einen Einstellungsfragebogen eben nicht drei, sondern *ein* inhaltliches, kognitiv verarbeitetes Evaluationsmuster abbilden. Eagly und Chaiken, deren ältere Arbeiten zur Begründung dieser drei Einstellungskomponenten häufig zitiert werden, schreiben selbst dazu: „Attitudes may indeed reflect cognitive, affective, and behavioral experiences with attitude objects, but there is little justification for claiming that these experiences produce three separable and omnipresent components of evaluative tendencies“ (Eagly & Chaiken, 2007, S. 589–590). Daher gehen wir bei der Entwicklung der PREIS von *einem* inhaltlichen evaluativen Muster aus. Bei der Formulierung und Auswahl der Items werden die drei Komponenten aber berücksichtigt. Andere übergeordnete Faktoren wie die Konsistenz und Stärke von Einstellungen sind theoretisch plausibel (siehe Kapitel 3 in Bohner & Wänke, 2014), aber aus methodischen Gründen nicht Gegenstand der vorliegenden Instrumentenentwicklung.

### *Problematische Itemformulierungen*

*Unklarer Einstellungsgegenstand.* Die Erfassung der Einstellungen zu Inklusion mittels Selbstbericht ist nur möglich, wenn den befragten Personen das Einstellungsobjekt bekannt ist (Schwarz & Bohner, 2001). Bisher werden den Befragten Fragebögen vorgelegt, ohne die zugrundeliegenden Konzepte von schulischer Inklusion zu explizieren. Es dürften aber große Unterschiede im Antwortverhalten resultieren, wenn die Befragten völlig unterschiedliche Verständnisse des Konstrukts haben. Das dürfte zum Beispiel zwischen verschiedenen pädagogischen Professionen, Lehrkräften verschiedener Schulformen und erst recht zwischen Menschen innerhalb und außerhalb pädagogischer Professionen der Fall sein. De facto bewerten dann verschiedene Befragte bei der Bearbeitung des Fragebogens *unterschiedliche* Einstellungs-

objekte, weshalb ihre Angaben nicht miteinander vergleichbar sind. Der PREIS ist daher eine präzise Instruktion zum inklusiven Schulsystem vorangestellt, damit alle Befragten das identische Einstellungsobjekt bewerten.

*Mangelnde Passung zum Einstellungsgegenstand.* Zwar gibt es keinen Konsens darüber, wie genau „Inklusion“, „schulische Inklusion“ oder „inklusives Schulsystem“ definiert werden können (Göransson & Nilholm, 2014; Nilholm & Göransson, 2016; Piezunka, Schaffus & Grosche, 2017), weitgehend konsensfähig ist aber die „zentrale Abkehr von der Zwei-Gruppen-Theorie“ (Grosche, 2015, S. 29; Dekategorisierung von Kindern z.B. mit / ohne Behinderung oder sonderpädagogischen Förderbedarf bzw. normal / besonders). Will man die Einstellungen zu einem solchen Schulsystem messen, müssen die Items in einer entsprechenden Einstellungsskala frei von ebendiesen Kategorisierungen sein. Ein wesentlicher Aspekt der Konstruktvalidität ist daher der Verzicht auf die Terminologie der Zwei-Gruppen-Theorie in den Items. In den aktuellen und etablierten Instrumenten ist dies nicht der Fall: „Wenn Kinder mit einer geistigen Entwicklungsverzögerung eine Regelklasse besuchen, dann leidet die Qualität des Unterrichts für die Kinder ohne Behinderung [Hervorh. d. Verf.]“ (Bosse & Spörer, 2014) oder „Kinder mit besonderen Bedürfnissen würden in einer inklusiven Schulklasse von den anderen Kindern gut behandelt werden [Hervorh. d. Verf.]“ (Seifried, 2015). Bei der Formulierung der Items haben wir daher auf Kategorisierungen verzichtet und andere Konstrukte möglichst herausgehalten.

*Priming.* Zusätzlich ist durch die Aktivierung solchen Gruppendenkens und der transportierten Annahmen über vermeintlich homogene (Sub-)Gruppen von Schülerinnen und Schülern über Itemformulierungen mit einem direkten Priming-Effekt (Bargh & Chartrand, 2014) zu rechnen. Aktiviert man durch die aktive Nutzung von Begrifflichkeit

ten der Zwei-Gruppen-Theorie Stereotype und das Denken in komplementären Kategorien (Kinder *mit* vs. Kinder *ohne* Behinderung, *normal* vs. *unnormal*) ist von einer unerwünschten negativen Beeinflussung der berichteten Einstellungen auszugehen. Eine solche Beeinflussung durch das Erhebungsinstrument selbst vermeidet die PREIS durch die Orientierung an einem nicht-kategorisierenden Schulsystem, in dem *alle* Schülerinnen und Schüler ohne (vermeintlich oder tatsächlich) stigmatisierende Zuschreibungen gefördert werden.

### *Mangelnde Abgrenzung von anderen Konstrukten*

Einige der vorliegenden Skalen enthalten Items, die eine Vermischung mit anderen Konstrukten verursachen, z.B. Selbstwirksamkeitserwartungen bezüglich inklusiven Unterrichts oder Einstellungen gegenüber Kindern mit Behinderungen. Erstens enthalten viele Instrumente Items, die Selbstwirksamkeitserwartungen bezüglich des Unterrichts in inklusiven Lernsettings oder bestimmter Lerngruppen erfassen: „Jeder Unterrichtsinhalt kann so aufbereitet werden, dass sowohl Ruben als auch David angemessenen Lernzuwachs erhalten [zwei Schüler mit stark unterschiedlichen Leistungen, Anm. d. Verf.]“ (Kopp, 2009) oder „Ich weiß, wie ich den spezifischen Bedürfnissen von Schülerinnen und Schülern mit Behinderungen in meinem Klassenraum begegnen kann“ (Przibilla, Lauterbach, Boshold, Linderkamp & Krezmien, 2016). Selbstverständlich sind Selbstwirksamkeitserwartungen von Lehrkräften relevant für das Unterrichtshandeln, aber es handelt sich dabei nicht um Einstellungen zu Inklusion. Dies ist auch daran zu erkennen, dass Einstellungen und Selbstwirksamkeit im Allgemeinen und im speziellen Fall von Einstellungen zu schulischer Integration in empirischen Studien zwar mittel bis hoch korreliert, aber doch distinkt sind (Bosse & Spörer, 2014:  $r = .64-.74$ ; Hellmich & Görel, 2014:  $r = .42-.56$ ). Diese Items setzen außerdem Wissen über

und Erfahrungen mit Lernprozessen und Unterrichtsabläufen voraus, die zumindest bei Menschen außerhalb der pädagogischen Profession nicht vorausgesetzt werden können.

Weiterhin beziehen sich viele Items der vorliegenden Instrumente weniger auf *Inklusion* als auf *das besondere Kind*, wie z.B. in der Aussage „Es ist wahrscheinlich, dass der/die Schüler/in mit besonderen Bedürfnissen in einer Regelklasse Verhaltensstörungen zeigen wird“ (ORI; Antonak & Larrivee, 1995; Übersetzung ins Deutsche von Benoit & Bless, 2014). Das ist in vielen Fällen gewünscht und legitim, und diese Einstellungen sind generell relevant für soziale Interaktionen im inklusiven Unterricht, aber es handelt sich unseres Erachtens nicht um Einstellungen zu Inklusion, sondern zu einer Gruppe von Kindern. Die PREIS zielt daher ausschließlich auf die Messung der Einstellung zum inklusiven Schulsystem ab.

### *Adressatengerichtetheit*

Es ist festzustellen, dass sich die durchaus umfassenden Forschungsaktivitäten im Bereich der Messung der Einstellungen zu Inklusion überwiegend auf die Einstellungen von (angehenden) Lehrkräften konzentrieren (Ruberg & Porsch, 2017; Schwab, Tretter & Gebhardt, 2013). Die Umsetzung von Inklusion im Schulsystem ist aber eine Reform, die sehr weitreichende Konsequenzen hat und keinesfalls nur Lehrerinnen und Lehrer, sondern auch Eltern oder das Schulumfeld sowie alle Bürgerinnen und Bürger (spätestens bei Wahlen) betrifft. Daher sind die Einstellungen von Nicht-Lehrkräften ebenso relevant für die Entwicklung eines inklusiven Schulsystems, bisher aber kaum erforscht. So resümieren Schwab, Tretter und Gebhardt: „Hinsichtlich der Forschung sollte die Einstellung der Allgemeinbevölkerung stärker in den Fokus treten“ (2013, S. 30). Zwar wurden in der Einstellungsforschung zum Thema „Schulische Inklusion“ auch verschiedene Gruppen von Lehrkräften miteinander oder Lehrkräfte mit anderen Personengruppen

verglichen (z.B. Schwab et al., 2013; Trumpa, Janz, Heyl & Seifried, 2014). Wir bezweifeln aber, dass eines der bis dato vorliegenden Instrumente dafür geeignet ist: Zum einen wegen der oben beschriebenen eher grundsätzlichen Zweifel an ihrer inhaltlichen Validität. Zum anderen, weil sie für die meisten Menschen ohne pädagogische Profession, teilweise sogar für Lehrkräfte ohne Erfahrungen mit inklusiven Lernsettings nicht sinnvoll einsetzbar sind, da einige oder sogar viele der Items konkretes Wissen über das Unterrichtshandeln von Lehrkräften und das Lernen von Schülerinnen und Schülern voraussetzen (siehe auch die Beispiele im vorherigen Abschnitt). Der Abruf von Wissen über und Erfahrungen mit dem Einstellungsobjekt sind aber zentrale Prozesse der Einstellungsbildung (Schwarz & Bohner, 2001). Die Items der PREIS sind daher so formuliert, dass sie professions- und erfahrungsunabhängig beantwortet werden können.

### *Antworttendenzen*

Selbstberichte als Messmethode für Einstellungen unterliegen einer Gruppe von systematischen Verzerrungen, den sogenannten Antworttendenzen, die weitgehend unumgänglich sind (aktueller Überblick: Wetzel, Böhnke & Brown, 2016), beispielsweise Akquieszenz (Zustimmungstendenz), Disakquieszenz (Ablehnungstendenz) und soziale Erwünschtheit (Paulhus, 2002; Uziel, 2010). Der Einfluss sozialer Erwünschtheit scheint insbesondere in Deutschland bei sozial sensiblen Themen (wie Inklusion) vergleichsweise groß zu sein (Leyser, Kapperman & Keller, 1994). Die positive Verzerrung der selbstberichteten Einstellung zu Inklusion durch die soziale Erwünschtheit wurde querschnittlich (Lui, Sin, Yang, Forlin & Ho, 2015) und experimentell (Lüke & Grosche, 2017c) nachgewiesen. Ausschließen lässt sich der Einfluss dieser Antworttendenzen bei Selbstberichten nicht, es stehen aber durchaus einige Methoden zur Reduktion oder Kontrolle ihres Einflusses zur Verfü-

gung. Diese Möglichkeiten bleiben im Bereich der Einstellungsforschung zu Inklusion bisher weitgehend ungenutzt und werden in der PREIS berücksichtigt.

### *Validierungsstudien*

Zwar wurde die Testgüte der meisten publizierten Einstellungsskalen empirisch überprüft. Aber die veröffentlichten Validierungsstudien lassen teilweise Fragen offen und die Replizierbarkeit ihrer Ergebnisse erscheint oft fraglich oder war nicht erfolgreich. Zudem basieren einige Studien auf sehr kleinen Stichproben, die Zweifel an der Anwendbarkeit der genutzten statistischen Verfahren und der Aussagekraft der Stichprobe aufkommen lassen. Nicht selten werden post-hoc bzw. explorativ – theoretisch kaum begründete – Subskalen gebildet oder vorher postulierte theoretisch begründete Modelle wegen fehlender Modellpassung verworfen, um Alternativmodelle vorzuschlagen. Die notwendige Überprüfung dieser Modelle an einer weiteren unabhängigen Stichprobe bleibt aber aus. In anderen Fällen werden Ergebnisse wie interne Konsistenz, Trennschärfe oder *fit indices* sehr großzügig interpretiert. Die Retest-Reliabilität wurde nach unserer Kenntnis noch bei keinem deutschsprachigen Verfahren überprüft. Daher legen wir hier eine ausführliche, konsekutive und strenge Prüfung der Testgüte der PREIS vor.

## **Konstruktion der Professionsunabhängigen Einstellungsskala zum Inklusiven Schulsystem (PREIS)**

Bei allen uns bekannten Instrumenten zur Messung von Einstellungen zu Inklusion (Lüke et al., 2017) sehen wir eines oder mehrere Probleme. Daher legten wir bei der grundsätzlichen Neuentwicklung der PREIS (1) aktuelle Theorien über die Struktur von Einstellungen zugrunde, (2) konstruierten die



Items konsistent zum Einstellungsgegenstand „inklusive Schulsystem“, (3) achteten sorgfältig auf möglichst geringe Vermischungen mit anderen Konstrukten, (4) formulierten die Items adressatengerecht und professionsübergreifend, (5) versuchten Verzerrungen durch Antworttendenzen systematisch zu reduzieren und (6) validierten unsere Skala iterativ an unabhängigen Stichproben.

Erstens verstehen wir unter Einstellungen eine „psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor.“ (Eagly & Chaiken, 1993, S. 1, 2007, S. 598). Nach der *Theory of Planned Behavior* (Ajzen, 1991) hängt das Auftreten eines bestimmten Verhaltens von den Einstellungen gegenüber diesem Verhalten, der wahrgenommenen sozialen Norm, der wahrgenommenen eigenen Verhaltenskontrolle sowie (teilweise daraus entstehenden) Verhaltensintention ab (Ajzen & Fishbein, 2005). Um diesen prädiktiven Zusammenhang zukünftig empirisch zu überprüfen, wird ein Instrument benötigt, das ohne die oben beschriebenen Probleme die Einstellungen zum inklusiven Schulsystem misst.

Zweitens legten wir den Fokus auf das „inklusive Schulsystem“ als Einstellungsgegenstand und verzichteten in den Itemformulierungen auf die Bezeichnung konkreter Gruppen von Schülerinnen und Schülern. Es sollen *alle* Kinder gemeint und negatives Priming durch das Instrument selbst verhindert werden. Um den Einstellungsgegenstand für alle Befragten möglichst konsistent zu halten, enthält die Instruktion der PREIS eine Definition des Einstellungsgegenstands. Dabei wird die Zuständigkeit der allgemeinen Schulen für alle Kinder als Kriterium genannt. ‚Alle Kinder‘ wird dann über die Aufzählung zahlreicher Heterogenitätsdimensionen noch einmal spezifiziert (Lücke & Grosche, 2017b).

Drittens versuchten wir verwandte Konstrukte, wie zum Beispiel Einstellungen gegenüber Menschen mit Behinderungen oder Selbstwirksamkeitserwartungen, aus den Items herauszuhalten.

Viertens formulierten wir alle Items so, dass auch Nicht-Pädagoginnen und -Pädagogen ohne eigene unterrichtliche Erfahrungen in inklusiven Schulklassen die Items problemlos beantworten konnten.

Fünftens versuchten wir den Einfluss von Antworttendenzen zu reduzieren bzw. kontrollierbar zu machen. Durch die Hinzunahme negativ formulierter Items, die durchaus negative Auswirkungen auf die theoretische Faktorenstruktur haben kann, werden Akquieszenz und Disakquieszenz als simple Antwortmuster offensichtlich erkennbar. Wesentlich schwieriger ist die Reduzierung sozial erwünschter Antworten. Wir haben versucht den Einstellungsgegenstand in den Items als recht abstraktes, distales, politisches Thema darzustellen, um alle denkbaren Positionen erkennbar zu legitimieren. Wir versuchen so den Einfluss sozialer Erwünschtheit zu reduzieren; wirkliche Lösungen erwarten wir aber erst durch den Einbezug von mehr als einer Methode zur Erfassung der Einstellungen, beispielsweise durch die Kombination expliziter und impliziter Methoden (Lücke & Grosche, 2017a). Zur statistischen Kontrolle der sozialen Erwünschtheit haben wir in allen hier berichteten Studien die wahrgenommene soziale Norm bezüglich Einstellungen zu Inklusion und in einer Studie die soziale Erwünschtheit umfassend erfasst.

Sechstens beschreiben wir replizierbar die Entwicklung und Validierung der PREIS. Die Itemformulierung erfolgte mittels der *Critical Incident Technique* (Flanagan, 1954): In Gesprächen mit dem Erstautor versetzten sich verschiedene Diskussions Teilnehmerinnen und -teilnehmer in die Situation, eine positive oder negative Einstellung gegenüber dem inklusiven Schulsystem zu haben, und formulierten daraufhin besondere Ereignisse (*Critical Incidents*), an denen man diese positive oder negative Einstellung besonders gut erkennen könne (Bühner, 2011). Dadurch wurden im ersten Schritt 39 mögliche Items formuliert. Im zweiten Schritt befragten wir fünf Expertinnen und Experten (Forscherinnen und For-

scher sowie Lehrkräfte) zur Inhaltsvalidität und Verständlichkeit der 39 Items. Daraufhin wurden mehrere Items überarbeitet und 9 Items entfernt. Im dritten Schritt wurden die resultierenden 30 Items einem *thinking aloud*-Test ( $N = 5$ ) unterzogen. Items, die aus Sicht der Antwortenden kompliziert waren oder nicht wie intendiert verstanden wurden, wurden entfernt. Es verblieb ein erster Itempool von 20 Items. Die weiteren im Folgenden beschriebenen Schritte bestanden aus vier Studien, in denen wir die Itemgüte und die Passung eines einfaktoriel- len Messmodells (Studien 1, 2 und 4), die mögliche Verzerrung durch soziale Erwünschtheit (Studie 2) und die Retest-Reliabilität (Studie 3) empirisch testeten. Um die professionsunabhängige Einsetzbarkeit der PREIS zu prüfen, verwendeten wir möglichst unterschiedliche Stichproben, die wir aus einem mittelgroßen Unternehmen (Studie 1), der universitären Lehrerbildung (Studie 2), Schulkollegien (Studie 3) und den Auszubildenden eines Berufsförderungswerks (Studie 4) rekrutierten.

## Generelle Methode

Die vollständige Instruktion der PREIS, wie sie im Anhang zu finden ist, wurde in allen Studien unverändert verwendet. In Studie 1 wurde der erste Itempool mit 20 Items verwendet. Aus unten diskutierten Gründen wurden sechs Items aus dem Instrument entfernt. In den Studien 2, 3 und 4 wurden immer dieselben 14 Items verwendet. Diese 14 Items sind in Tabelle 1 dokumentiert. Als Antwortformat wurde in allen Studien eine fünfstufige Likert-Skala (0 bis 4) mit den semantischen Ankern *stimme nicht zu* und *stimme zu* verwendet.

Im Anschluss an den Einstellungsfragebogen wurden in allen Studien folgende zusätzliche Angaben erfragt: Geschlecht, Alter, eigene Kinder, eigene schulpflichtige Kinder, regelmäßiger (mindestens monatlicher) Kontakt zu Menschen mit Behinderung – beispielsweise in der Familie oder im Freun-

deskreis, allgemeine politische Einstellung (10-stufige Likert-Skala; *links* = 0, *rechts* = 9), zur Prüfung der sozialen Erwünschtheit die wahrgenommene Einstellung zu Inklusion der befragenden Institution (5-stufige Likert-Skala; *absolut dagegen* = 0, *absolut dafür* = 4). Als Anreiz wurden unter allen Teilnehmerinnen und Teilnehmern, die den Fragebogen abgeschlossen haben, zwei Einkaufsgutscheine im Wert von 50€ verlost.

Das Datenmanagement und die Analysen erfolgten mit R 3.3.2 (R Core Team, 2016) in RStudio 1.0.136 (RStudio Team, 2016) unter Anwendung der Pakete *psych* 1.7.5 (Revelle, 2017), *car* 2.1-4 (Fox & Weisberg, 2011) und *lavaan* 0.5-23.1097 (Rosseel, 2012).

Die Prüfung des Messmodells erfolgt über eine Konfirmatorische Faktorenanalyse (CFA) mit *Robust Maximum Likelihood*-Schätzer, um dem Antwortformat gerecht zu werden. Zur Prüfung der Modellpassung werden *data-model fit indices* aller Klassen (*absolute*, *parsimonious*, *incremental*) herangezogen (Jackson, Gillaspay & Purc-Stephenson, 2009). Nach Hu und Bentler (1999) wurden dabei als Kriterien für eine gute Modellpassung  $CFI$  &  $TLI \geq .95$ ,  $SRMR \leq .08$  sowie  $RMSEA \leq .05$  angewendet.

## Studie 1

### Ziele

In der ersten Studie sollte das Instrument im Feld erprobt, dabei problematische Items identifiziert und, sofern inhaltlich vertretbar, die Anzahl der Items reduziert werden. Dazu wurde der Fragebogen einer größeren Stichprobe von Menschen vorgelegt, die in keinem pädagogischen Handlungsfeld arbeiten. Da diese Personen über weniger Vorwissen zum inklusiven Schulsystem verfügen dürften, konnte hiermit die Einsetzbarkeit der Instruktion und der Items unter anspruchsvollen Bedingungen erprobt werden. Items mit großer inhaltlicher Passung (hoher Vali-

dität), sollten im Zweifelsfall auf Kosten der statistischen Passung im Fragebogen verbleiben.

## Methoden

### Stichprobe

Es wurden 167 Mitarbeiterinnen und Mitarbeiter (62% weiblich) eines mittelgroßen Unternehmens im Alter von durchschnittlich 34.90 Jahren ( $SD = 12.07$ ) für die Teilnahme an der Studie gewonnen. Keine der befragten Personen arbeitete in einem pädagogischen Beruf. Etwa die Hälfte (47%) von ihnen hatte eigene Kinder, 32% Kinder, die bereits im schulpflichtigen Alter waren. Dreißig Prozent der Teilnehmenden hatten regelmäßigen Kontakt zu einem Menschen mit Behinderung. Die durchschnittliche politische Einstellung lag bei  $M = 3.58$  ( $SD = 1.51$ ) und ist minimal „links“ des theoretischen Skalenmittelwerts zu interpretieren. Die durchschnittlich wahrgenommene Einstellung der befragenden Institution war mit  $M = 2.74$  ( $SD = 0.91$ ) eher positiv.

### Instrumente & Prozedere

Die 20 Items des eingangs beschriebenen Itempools sowie die zusätzlichen Angaben zur Person wurden als Onlinefragebogen (LimeSurvey 2.65.1, 2016) präsentiert. Die Teilnehmenden wurden vor und nach der Befragung zu Rückmeldungen bezüglich des Instruments und einzelner Items ermutigt. Dazu wurden ein offenes Antwortfeld sowie die E-Mail-Adresse des Projektleiters angeboten. Die Befragten wurden über den firmeninternen E-Mail-Verteiler angeschrieben und um Teilnahme gebeten. Die Datenerhebung erfolgte in einem Zeitraum von sechs Wochen. Die durchschnittliche Bearbeitungszeit lag bei  $Md = 6$  Minuten (Interquartilsabstand  $[IQR] = 3$ ).

## Ergebnisse

Zwei Items, die im Rahmen der Pretests vereinzelt auch schon kritisiert wurden, aber zunächst im Itempool verblieben („Ich bin bereit mich für den Erhalt unseres jetzigen Schulsystems zu engagieren.“ und „Etwas in mir sagt, dass ein inklusives Schulsystem keine sehr gute Idee ist.“), wurden vermehrt ausgelassen, kommentiert und in direktem Feedback an die Studienleitung kritisiert. Beide Items wurden daraufhin endgültig entfernt, um zukünftig fehlende Werte zu vermeiden.

Die 18 verbleibenden Items wurden ohne fehlende Werte von allen Teilnehmerinnen und Teilnehmern beantwortet. Keines der Items zeigte einen Boden- oder Deckeneffekt ( $Min_M = 1.59$ ,  $Max_M = 2.74$ ,  $M_M = 2.13$ ,  $SD_M = 0.40$ ). Die Überprüfung des Einfaktormodells mittels CFA ergab keine gute Modellpassung:  $\chi^2 = 348.74$ ,  $df = 135$ ,  $p < .001$ ,  $CFI = .89$ ,  $TLI = .89$ ,  $SRMR = .05$ ,  $RMSEA = .10$ . Wegen zu stark korrelierter Faktoren konnte ein Dreifaktormodell mit den drei Einstellungskomponenten nicht geprüft werden. Anhand der Itemkennwerte und *mod indices* wurden vier besonders auffällige Items identifiziert: „Ich halte ein inklusives Schulsystem in Deutschland für nicht umsetzbar.“, „Ich halte ein inklusives Schulsystem in meinem Ort für nicht umsetzbar.“, „Ich denke, ein inklusives Schulsystem wäre nicht in der Lage, alle Kinder zu ihren bestmöglichen Leistungen zu bringen.“ und „Ich denke, dass einige Kinder in einem inklusiven Schulsystem in der Masse untergehen und nicht passend gefördert würden.“.

Nach dem Entfernen dieser vier Items verbesserte sich die Modellpassung:  $\chi^2 = 156.84$ ,  $df = 77$ ,  $p < .001$ ,  $CFI = .95$ ,  $TLI = .94$ ,  $SRMR = .04$ ,  $RMSEA = .08$ . Obwohl diese nicht nach allen angelegten Kriterien gut ist, wurde das Messmodell als ausreichend passend akzeptiert, um es in einer weiteren Studie an einer größeren Stichprobe zu validieren. Das dreifaktorielle Modell erschien wegen zu hoher Faktorinterkorrelationen weiterhin unplausibel.



## Diskussion

In dieser Studie sollte das Instrument erstmals einer größeren Stichprobe vorgelegt werden. Abgesehen von zwei Items (die dementsprechend aus dem Itempool entfernt wurden) gab es bei den Befragten – allesamt keine pädagogischen Fachkräfte – keine Verständnisschwierigkeiten. Anhand der Daten wurden vier weitere Items zur Kürzung des Instruments identifiziert. In allen vier Fällen sind die Gründe für die mangelnde Passung inhaltlich erkennbar und auf Fehler bei der Itemkonstruktion zurückzuführen: Die ersten beiden Items zielen in ihrer Formulierung auf die Umsetzbarkeit eines inklusiven Schulsystems – diese Einschätzung ist aber offensichtlich unabhängig von der eigenen Einstellung einer Person. Die beiden anderen Items zielten auf die Förderung *aller* Kinder, wurden aber ungünstig formuliert, da die Aussagen in ihrer Absolutheit auf kein denkbares Schulsystem zutreffen dürften. Diese Einschätzung dürfte also weniger eine spezifische Einstellung zum inklusiven Schulsystem als vielmehr eine allgemeine Einschätzung von Schule beinhalten. Die problematischen Items wurden entfernt.

## Studie 2

### Ziele

Die auf 14 Items gekürzte Fragebogenversion wurde anschließend bei einer neuen und größeren Stichprobe eingesetzt. Dabei sollten die Itemkennwerte und die faktorielle Struktur des überarbeiteten Fragebogens geprüft werden, gegebenenfalls weitere unpassende Items identifiziert und aus dem Fragebogen gekürzt werden.

Basierend auf den Ergebnissen von Studie 1 erwarten wir, dass die Items keine Boden- oder Deckeneffekte zeigen (H1: Die mittleren Itemschwierigkeiten liegen zwischen 1.00 und 3.00) und trennscharf sind (H2: Die korrigierten Skala-Item-Korrelationen sind  $\geq .30$ ). Weiterhin sollte die PREIS

das eindimensionale Konstrukt „Einstellungen zum inklusiven Schulsystem“ abbilden (H3: Ein einfaktorielles Messmodell zeigt gute Passung). Ein Ziel der Konstruktion der PREIS war es auch den Einfluss sozialer Erwünschtheit durch eine eher distale Verortung des Einstellungsgegenstandes gering zu halten. Die Einstellungen – gemessen mit der PREIS – sollten daher gering mit der Tendenz zu sozial erwünschten Antworten korrelieren (H4: Korrelation zwischen PREIS und einer Skala der sozialen Erwünschtheit ist  $r < .30$ ).

### Methode

#### Stichprobe

Es nahmen 305 Lehramtsstudierende (86% weiblich) im Alter von durchschnittlich 21.60 Jahren ( $SD = 2.66$ ) an der Studie teil. Sie befanden sich zum Befragungszeitpunkt im Mittel ( $Md$ ) im dritten Fachsemester ( $IQR = 3$ ). Zwei Prozent von ihnen hat eigene Kinder, 1% Kinder im schulpflichtigen Alter. Neununddreißig Prozent der Teilnehmenden hatten regelmäßigen Kontakt zu einem Menschen mit Behinderung. Die mittlere politische Einstellung lag bei  $M = 3.03$  ( $SD = 1.38$ ) und ist etwas „links“ der theoretischen Mitte zu interpretieren. Die wahrgenommene Einstellung der befragenden Institution ist mit  $M = 2.90$  ( $SD = 0.84$ ) als eher positiv zu beschreiben.

#### Instrumente & Prozedere

Die Studierenden wurden in mehreren Lehrveranstaltungen einer nordrhein-westfälischen Universität um ihre Teilnahme gebeten. Es wurde jedoch klargestellt, dass die Teilnahme an der Studie freiwillig und vollkommen unabhängig von der Lehrveranstaltung war. Die Studierenden beantworteten die Onlineversion des Fragebogens unter Nutzung eines Pseudonyms, das weitere Untersuchungen im Längsschnitt erlaubt. Zwischen PREIS und den soziodemographischen Angaben beantworteten die

Teilnehmerinnen und Teilnehmer zusätzlich die „Kurzskala Soziale Erwünschtheit-Gamma“ (KSE-G; Kemper, Beierlein, Bensch, Kovaleva & Rammstedt, 2012). Diese basiert auf dem von Paulhus (2002, 2007) vorgeschlagenen Modell sozialer Erwünschtheit und erfasst die Übertreibung positiver Qualitäten ( $PQ^+$ ,  $\omega = .71$ ) sowie die Untertreibung negativer Qualitäten ( $NQ^-$ ,  $\omega = .78$ ) mit insgesamt sechs Items. Die Korrelation zwischen dem PREIS-Score und der wahrgenommenen Einstellung der befragenden Institution liegt bei  $r = .33$ ,  $p < .001$ . Die durchschnittliche Bearbeitungszeit des gesamten Fragebogens lag bei  $Md = 8$  Minuten ( $IQR = 3$ ).

## Ergebnisse

Im Durchschnitt erreichen die Teilnehmerinnen und Teilnehmer einen PREIS-Rohwert von 37.17 ( $SD = 11.22$ ). Die Itemschwierigkeiten (H1) bewegen sich im mittleren Bereich der Antwortskala ( $Min_M = 1.98$ ,  $Max_M = 3.18$ ,  $M_M = 2.66$ ,  $SD_M = 0.37$ ) und die Varianzen sind vergleichbar ( $M_{SD} = 1.12$ ,  $SD_{SD} = 0.11$ ). Die korrigierten Skala-Item-Korrelationen (H2) liegen im mittleren bis hohen Bereich ( $M_{rit} = .69$ ,  $SD_{rit} = .10$ ). Die Passung zum einfaktoriellem Modell (H3) ist gut:  $\chi^2 = 135.76$ ,  $df = 77$ ,  $p < .001$ ,  $CFI = .97$ ,  $TLI = .97$ ,  $SRMR = .03$ ,  $RMSEA = .05$ . Es besteht keine signifikante Korrelation zwischen der PREIS und der KSE-G (H4;  $r = .02$ ,  $p > .05$ ). Die interne Konsistenz der Skala beträgt  $\alpha = .93$ .

## Diskussion

Ziel dieser Studie war es, die aus Studie 1 abgeleitete Version der Skala an einer neuen und größeren Stichprobe konfirmatorisch zu prüfen und ggf. weitere zu leichte, zu schwere oder zu wenig trennscharfe Items zu identifizieren. Die Itemschwierigkeiten und -varianzen sind zufriedenstellend. Insgesamt erzielten die Lehramtsstudierenden eher hohe Skalenrohwerter, sodass insbesondere nach unten weiter differenziert werden

kann. Die Trennschärfen bzw. standardisierten Faktorladungen (siehe Tabelle 1) sind ebenfalls zufriedenstellend. Die Passung zum theoretisch erwarteten einfaktoriellem Modell ist gut.

Für diese Stichprobe ist keine Korrelation zwischen PREIS und sozial erwünschter Antworttendenz nachweisbar. Wir werten dies als Hinweis darauf, dass unser Ansatz bei der Itemformulierung zur Reduzierung dieses Einflusses erfolgreich gewesen sein könnte. Da es sich bei der KSE-G aber um eine Kurzskala mit gerade ausreichender Testgüte handelt, müssen diese Befunde mit einem anderen Instrument zur Erfassung der sozial erwünschten Antworttendenz repliziert werden.

Die interne Konsistenz von  $\alpha = .93$  ist hoch, allerdings ist Cronbachs  $\alpha$  aus verschiedenen Gründen kein geeignetes Maß für die Reliabilität eines Instruments (Sijtsma, 2009). Da eine gewisse Stabilität der Einstellungen gegenüber dem inklusiven Schulsystem angenommen werden darf, prüfen wir in Studie 3 die Reliabilität der PREIS-Scores mit der Retest-Methode.

## Studie 3

### Ziele

Basierend auf gängigen Testbewertungsrichtlinien, wie zum Beispiel dem COTAN System (Evers, 2001), bewerten wir eine Retest-Reliabilität  $> .70$  für Auswertungen auf Gruppenebene und  $> .80$  für die Interpretation auf individueller Ebene als gut. Wir erwarten, dass die PREIS eine Test-Retest-Korrelation zeigt, die hoch genug ist, um nicht nur auf Gruppenebene, sondern auch auf Individuumsebene reliable Messungen zu ermöglichen (H1: Die Test-Retest-Korrelation ist  $r \geq .80$ ).

## Methodie

### Stichprobe

Es wurden 57 Gymnasiallehrkräfte (61% weiblich) zweier Gymnasien im Alter von durchschnittlich 45.25 Jahren ( $SD = 13.14$ ) für die Teilnahme an der Studie gewonnen. Rund 74% von ihnen hatten eigene Kinder, 56% Kinder, die bereits im schulpflichtigen Alter waren. Einunddreißig Prozent der Teilnehmenden hatten regelmäßigen Kontakt zu einem Menschen mit Behinderung. Die wahrgenommene Einstellung der befragenden Institution war mit  $M = 2.65$  ( $SD = 0.90$ ) als eher positiv zu beschreiben.

### Instrumente & Prozedere

Wir baten die Lehrkräfte zweier Gymnasien in Nordrhein-Westfalen um die Teilnahme an der Studie. Auf Wunsch der Schulleitungen wurden die Fragebögen als Papierversionen eingesetzt und die Frage nach der allgemeinen politischen Einstellung im Vorfeld entfernt. Zum ersten Messzeitpunkt (t1) füllten 83 Lehrkräfte den Fragebogen vollständig aus. Zum zweiten Messzeitpunkt (t2) nach vier Wochen füllten noch 57 von ihnen den Fragebogen erneut vollständig aus. Erst im Anschluss wurden sie über den genauen Zweck der Studie informiert. Die Korrelation zwischen dem PREIS-Score und der wahrgenommenen Einstellung der befragenden Institution liegt bei  $r = .26$ ,  $p > .05$ .

### Ergebnisse

Der durchschnittliche PREIS-Score der 57 Teilnehmenden war bei der ersten Befragung  $M_{t1} = 22.09$  ( $SD_{t1} = 15.20$ ) und bei der zweiten Befragung  $M_{t2} = 21.00$  ( $SD_{t2} = 14.47$ ). Die Test-Retest-Korrelation (H1) nach vier Wochen betrug  $r = .96$ ,  $t(55) = 26.04$ ,  $p < .001$ .

## Diskussion

Die vierwöchige Retest-Reliabilität von  $r = .96$  in dieser Stichprobe ist als ein deutlicher Hinweis auf die hohe Zuverlässigkeit der Messungen mit der PREIS zu bewerten. Damit eignet sich die PREIS für Auswertungen sowohl auf Gruppen- als auch auf Individualebene. Hier ist insbesondere wichtig zu betonen, dass nach unserer Kenntnis bisher noch für keines der anderen deutschsprachigen Instrumente eine Untersuchung der Retest-Reliabilität vorliegt. Aufgrund der kleinen und eher homogenen Stichprobe sind die Ergebnisse dieser Studie allerdings nur vorsichtig zu interpretieren. Weiterhin haben nur zwei Drittel der Befragten an beiden Erhebungen teilgenommen. Zwar fanden wir keine Hinweise auf systematischen Dropout. Aber eine Verzerrung im Hinblick auf das Engagement und Interesse für das Thema ist nicht auszuschließen.

## Studie 4

### Ziele

Ziele der vierten Studie waren die erneute Prüfung der Itemgüte und der Faktorenstruktur an einer weiteren Stichprobe von Personen ohne pädagogischen Hintergrund, um die grundsätzliche Professionsunabhängigkeit der PREIS zu replizieren. Basierend auf den Ergebnissen der Studien 1 und 2 erwarten wir, dass die Items keine Boden- oder Deckeneffekte zeigen (H1: Die mittleren Itemschwierigkeiten liegen zwischen 1.00 und 3.00) und trennscharf sind (H2: Die korrigierten Skala-Item-Korrelationen sind  $\geq .30$ ). Weiterhin sollte die PREIS das ein-dimensionale Konstrukt „Einstellungen zum inklusiven Schulsystem“ abbilden (H3: Ein einfaktorielles Messmodell zeigt gute Passung).

## Methode

### Stichprobe

Es wurden 208 Auszubildende zur Altenpflegerin bzw. zum Altenpfleger (75% weiblich) eines nordrhein-westfälischen Berufsförderungswerks im Alter von durchschnittlich 25.85 Jahren ( $SD = 9.14$ ) für die Teilnahme an der Studie gewonnen. Keine der befragten Personen arbeitete in einem pädagogischen Beruf. Knapp ein Drittel (30%) von ihnen hatte eigene Kinder, 22% Kinder, die bereits im schulpflichtigen Alter waren. Fünfundvierzig Prozent der Teilnehmenden gaben an regelmäßigen Kontakt mit einem Menschen mit Behinderung in ihrer Familie oder ihrem Freundeskreis zu haben. Im Rahmen ihrer beruflichen Tätigkeit traf dies auf 88 Prozent der Befragten zu. Die durchschnittlich wahrgenommene Einstellung der befragenden Institution war mit  $M = 2.55$  ( $SD = 1.15$ ) eher positiv.

### Instrumente & Prozedere

Die Auszubildenden wurden über die Lehrkräfte einer nordrhein-westfälischen Berufsförderungswerk um ihre Teilnahme gebeten. Es wurde jedoch klargestellt, dass die Teilnahme an der Studie freiwillig war. Die Auszubildenden beantworteten die Papierversion des Fragebogens. Die Korrelation zwischen dem PREIS-Score und der wahrgenommenen Einstellung der befragenden Institution liegt bei  $r = .20$ ,  $p < .05$ . Drei der ursprünglich 211 ausgefüllten Fragebogen wurden aus der Stichprobe entfernt, weil die Antwortenden nicht-interpretierbare Muster angekreuzt hatten (z.B. ausschließlich *stimme zu* [4] bei allen – auch negativ gepolten – Items). In 17 Datensätzen traten unsystematisch fehlende Werte auf (14 Muster; 1-2 fehlende Werte pro Fall). Die fehlenden Werte wurden im Rahmen der CFA mit *full information maximum likelihood*-Schätzer ersetzt.

## Ergebnisse

Im Durchschnitt erreichen die Teilnehmerinnen und Teilnehmer einen PREIS-Rohwert von 35.74 ( $SD = 14.46$ ). Die Itemschwierigkeiten (H1) bewegen sich im mittleren Bereich der Antwortskala ( $Min_M = 2.32$ ,  $Max_M = 2.83$ ,  $M_M = 2.56$ ,  $SD_M = 0.17$ ) und die Varianzen sind vergleichbar ( $M_{SD} = 1.31$ ,  $SD_{SD} = 0.09$ ). Die korrigierten Skala-Item-Korrelationen (H2) liegen im mittleren bis hohen Bereich ( $M_{rit} = .74$ ,  $SD_{rit} = .04$ ). Die Passung zum einfaktoriellem Modell (H3) ist gut:  $\chi^2 = 148.85$ ,  $df = 77$ ,  $p < .001$ ,  $CFI = .96$ ,  $TLI = .96$ ,  $SRMR = .03$ ,  $RMSEA = .07$ .

## Diskussion

Ziel dieser Studie war die Replikation der Itemgüte und der konfirmatorischen Passung der PREIS in einer weiteren Stichprobe von Personen, die keiner pädagogischen Profession angehören. Die Itemschwierigkeiten und -varianzen sind zufriedenstellend und weichen nur geringfügig von den Befunden der vorherigen Studien ab. Insgesamt erzielen die Nicht-Pädagoginnen und -Pädagogen etwas niedrigere Skalenrohwerte als die Lehramtsstudierenden aus Studie 2. Die Trennschärfen bzw. standardisierten Faktorladungen sind ebenfalls zufriedenstellend. Die Passung zum theoretisch erwarteten einfaktoriellem Modell ist gut. Insgesamt sprechen die Befunde dafür, dass die PREIS auch bei Personen ohne pädagogische Profession das einfaktoriellem Konstrukt „Einstellungen zum inklusiven Schulsystem“ misst.

Es fällt auf, dass die korrigierten Skala-Item-Korrelationen und Itemschwierigkeiten in dieser Stichprobe etwas konsistenter sind als in der Stichprobe der Lehramtsstudierenden (Studie 2). Eine mögliche und theoretisch naheliegende Erklärung hierfür wäre, dass Personen mit einfacheren Repräsentationen des Einstellungsobjekts etwas globaler antworten, während Personen mit komplexeren Repräsentationen des Einstellungs-

objekts, stärker zwischen den Items differenzieren.

## Gesamtdiskussion

Die bisher verfügbaren Einstellungsskalen sind aus verschiedenen methodischen und inhaltlichen Gründen nur eingeschränkt in der Lage die Einstellungen zum inklusiven Schulsystem valide zu erfassen. Ziel unserer Arbeit war die grundlegende Neuentwicklung eines Instrumentes zur Messung von selbstberichteten Einstellungen zum inklusiven Schulsystem, das (1) auf aktuellen Einstellungstheorien basiert, (2) klare Itemformulierungen und einen präzisen Einstellungsgegenstand hat, (3) möglichst gering mit anderen Konstrukten vermischt ist, (4) professions- bzw. adressatenübergreifend einsetzbar ist, (5) den Einfluss von Antworttendenzen reduziert und (6) an unabhängigen Stichproben validiert wurde.

In einem mehrstufigen Prozess wurden Items entwickelt, revidiert, umformuliert oder ausgeschlossen. Letztendlich können wir in Tabelle 1 nun die sorgfältig konstruierte und psychometrisch in vier unabhängigen Stichproben validierte „Professionsunabhängige Einstellungsskala zum Inklusiven Schulsystem“ (PREIS) vorlegen. Zusammenfassend ist die PREIS ein Instrument zur Erfassung von Einstellungen zum inklusiven Schulsystem, das (a) valide das einfaktorielle Konstrukt Einstellungen zum inklusiven Schulsystem misst, ohne professionsspezifisches Wissen vorauszusetzen; (b) auf Formulierungen im Sinne der Zwei-Gruppen-Theorie verzichtet; (c) inhaltliche Antworttendenzen reduziert und ansatzweise kontrollierbar macht; (d) reliable und stabile Ergebnisse hervorbringt; und (e) für alle Interessierten frei und offen unter einer CC-BY-SA-Lizenz verfügbar ist (Lüke & Grosche, 2017b).

Trotz dieser recht aufwändigen Validierungsstudien muss sich das Instrument in anderen Studien empirisch bewähren. Weitere Teilstudien zur Konstruktvalidität und zum

Einfluss der vorangestellten Definition sind daher in Vorbereitung. Unabhängige Replikationen aller (Teil-) Studien ermutigen wir ausdrücklich.

Eine wichtige Einschränkung der Verwendung der PREIS ergibt sich durch das von uns definierte Einstellungsobjekt. Wir haben unseren Items eine explizite Definition eines inklusiven Schulsystems, das niemanden ausschließt und ohne Kategorien wie Behinderung oder Förderbedarf auskommt, vorangestellt. Wenn die Einstellungen zu dieser Art von Schulsystem erfragt werden sollen und das hier dementsprechend weit definierte Inklusionsverständnis passend sind, stellt die PREIS nach unserer Kenntnis die bestmögliche Alternative dar. Sollen jedoch Einstellungen zu Menschen mit Behinderungen und ihrer Förderung in Regelklassen erfragt werden, so sollte auf andere Verfahren zurückgegriffen werden (z.B. Subskalen von Bosse & Spörer, 2014; Przi-billa et al., 2016; Schwab et al., 2013; Seifried, 2015), sofern die psychometrische Testgüte für die jeweilige Anwendung ausreichend ist.

Eine weitere Einschränkung, die sich nicht nur auf die PREIS, sondern generell auf Selbstberichte bezieht, ist die Tatsache, dass einige Antworttendenzen nicht vollständig vermeidbar sind. Zwar haben wir bei der Skalenkonstruktion einige Antworttendenzen kontrollierbar gemacht bzw. reduziert. Wirklich lösen lässt sich das Problem aber nur durch die Kombination mehrerer Erhebungsmethoden und die Kontrolle sozialer Normen. Wir verfolgen daher das Ziel Einstellungen zum inklusiven Schulsystem unter anderem auch mit impliziten Methoden zu erfassen. Einen ersten Vorschlag für ein solches Instrument haben wir an anderer Stelle evaluiert (Lüke & Grosche, 2017a) und ebenfalls offen und kostenfrei zur Verfügung gestellt (Lüke, 2017).

Das dringendste Desiderat der Einstellungsforschung zum Thema Inklusion stellt sicher die Prüfung des Zusammenhangs von Einstellungen und später gezeigtem Verhalten dar. Bislang existiert unseres Wissens



Tabelle 1: Professionsunabhängige Einstellungsskala zum inklusiven Schulsystem mit Itemschwierigkeit, -streuung, -Skala-Trennschärfe und Parameterschätzungen des einfaktoriellem Messmodells aus Studie 2

Nr.	Item	M	SD	$r_{it}$ (korr.)	CFA	
					Faktorladung	Residualvarianz
01	Ich vermute, dass ein inklusives Schulsystem den Unterricht für alle Kinder verbessern könnte.	2.41	1.07	.74	0.75	0.44
02	Ich finde, dass die Einführung eines inklusiven Schulsystems unsere Gesellschaft positiv beeinflussen würde.	3.04	0.99	.78	0.78	0.40
03	Ich finde, ein inklusives Schulsystem wäre gerechter als das jetzige Schulsystem.	2.62	1.11	.69	0.69	0.53
04	<i>Bei einem Volksentscheid würde ich wahrscheinlich gegen ein inklusives Schulsystem stimmen.</i>	2.80	1.34	.75	0.75	0.44
05	Ich schätze, dass ein inklusives Schulsystem für einen größeren Zusammenhalt unter den Kindern sorgen würde.	2.99	1.06	.61	0.60	0.64
06	Der Gedanke an ein inklusives Schulsystem stimmt mich fröhlich.	2.49	1.04	.81	0.81	0.34
07	<i>Ich finde es unrealistisch, Kinder mit sehr unterschiedlichen Schwierigkeiten und Bedürfnissen in einer Klasse unterrichten zu wollen.</i>	2.00	1.29	.66	0.66	0.57
08	Ein inklusives Schulsystem entspricht meinen moralischen Werten.	3.18	0.95	.68	0.68	0.54
09	<i>Ich meine, dass ein inklusives Schulsystem mit dem Leistungsprinzip unserer Gesellschaft nicht vereinbar ist.</i>	1.98	1.19	.39	0.38	0.86
10	Eine Partei, die sich gegen ein inklusives Schulsystem stellt, wäre für mich nicht wählbar.	2.61	1.21	.64	0.63	0.60
11	<i>Ich fühle mich unwohl, wenn ich mir ein inklusives Schulsystem vorstelle.</i>	3.09	1.06	.73	0.73	0.46
12	Ich würde viel dafür tun, dass mein eigenes Kind in einem inklusiven Schulsystem unterrichtet wird.	2.47	1.10	.75	0.75	0.43
13	Ich wäre bereit, mich an der Entwicklung eines inklusiven Schulsystems zu beteiligen.	2.85	1.14	.72	0.72	0.48
14	Es wäre schön, wenn zukünftig alle Kinder eine gemeinsame Schule in einem inklusiven Schulsystem besuchen würden.	2.65	1.10	.77	0.77	0.40

Anmerkung. Standardisierte Koeffizienten des einfaktoriellem Messmodells (unkorrelierte Fehler) mit guter Modellpassung ( $\chi^2 = 135.76$ ,  $df = 77$ ,  $p < .001$ ,  $CFI = .97$ ,  $TLI = .97$ ,  $SRMR = .03$ ,  $RMSEA = .05$ ). Negativ gepolte Items (kursiv) wurden für die Analysen rekodiert.

keine Studie hierzu, obwohl diese Prädiktion – zumindest im Kontext der Einstellungen von Lehrkräften – die zentrale Legitimation vieler Forscherinnen und Forscher ist.

## Anhang

### **Definition schulischer Inklusion aus der Instruktion der PREIS**

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 fühlen sich verantwortlich für die Förderung und Unterstützung aller Kinder – unabhängig von ihren Lernvoraussetzungen, Erfahrungen und Bedürfnissen. Sie bieten einen Unterricht an, der allen Kindern Lernmöglichkeiten bietet, und schließen niemanden davon aus.“

Damit sind ausdrücklich alle Kinder gemeint, also nicht nur diejenigen mit Behinderungen. Denken Sie auch an alle möglichen Ethnien und Religionen, schwierige Lebenslagen, Armut, Delinquenz, besondere soziale und emotionale Bedürfnisse, unterschiedliche Herkunftsfamilien sowie Talente und Begabungen.“

Bitte denken Sie bei der Beantwortung der Fragen nur an diese Definition eines inklusiven Schulsystems, auch wenn Sie persönlich eine ganz andere Definition verwenden würden.

## Literaturverzeichnis

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50, 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Ajzen, I. & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracín, B. T. Johnson & M. P. Zanna (Eds.), *The Handbook of Attitudes* (pp. 173-221). Mahwah, NJ: Erlbaum.
- Antonak, R. F. & Larrivee, B. (1995). Psychometric analysis and revision of the opinions relative to mainstreaming scale. *Exceptional Children*, 62, 139–149.
- Avramidis, E. & Norwich, B. (2002). Teachers' attitudes towards integration / inclusion: A review of the literature. *European Journal of Special Needs Education*, 17, 129–147. <https://doi.org/10.1080/08856250210129056>
- Bargh, J. A. & Chartrand, T. L. (2014). The mind in the middle. A practical guide to priming and automaticity research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (Second edition, pp. 311-344). New York, NY: Cambridge University Press.
- Benoit, V. & Bless, G. (2014). Erfassung der Einstellungen zur schulischen Integration von Lehrpersonen: Übersetzung und Erprobung der ORI und ATIES Skalen. *Zeitschrift für Bildungsforschung*, 4, 209–226. <https://doi.org/10.1007/s35834-014-0110-z>
- Bohner, G. & Wänke, M. (2014). *Attitudes and attitude change* (Social psychology). London, UK: Psychology Press.
- Bosse, S. & Spörer, N. (2014). Erfassung der Einstellung und der Selbstwirksamkeit von Lehramtsstudierenden zum inklusiven Unterricht. *Empirische Sonderpädagogik*, 6, 279–299.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (PS Psychologie, 3., aktualisierte und erw. Aufl.). München: Pearson.
- De Boer, A., Pijl, S. J. & Minnaert, A. E. M. G. (2011). Regular primary schoolteachers' attitudes towards inclusive education: A review of the literature. *International Journal of Inclusive Education*, 15, 331–353.
- De Boer, A., Timmerman, M., Pijl, S. J. & Minnaert, A. (2012). The psychometric evaluation of a questionnaire to measure attitudes towards inclusive education. *European Journal of Psychology of Education*, 27, 573–589. <https://doi.org/10.1007/s10212-011-0096-z>

- Eagly, A. H. & Chaiken, S. (1993). *The Psychology of Attitudes*. Fort Worth, TX: Harcourt.
- Eagly, A. H. & Chaiken, S. (2007). The advantages of an inclusive definition of attitude. *Social Cognition, 25*, 582–602.
- Eder, F., Altrichter, H. & Paseka, A. (Hrsg.). (2014). Einstellungen zum inklusiven Unterricht [Themenheft]. *Zeitschrift für Bildungsforschung, 4*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Evers, A. (2001). Improving test quality in the Netherlands. Results of 18 years of test ratings. *International Journal of Testing, 1*, 137–153. [https://doi.org/10.1207/S15327574IJT0102\\_3](https://doi.org/10.1207/S15327574IJT0102_3)
- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 51*, 327–358. <https://doi.org/10.1037/h0061470>
- Fox, J. & Weisberg, S. (2011). *An R Companion to applied regression* (Second Edition). Thousand Oaks, CA: Sage.
- 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*, 265–280. <https://doi.org/10.1080/08856257.2014.933545>
- Grosche, M. (2015). Was ist Inklusion? In P. Kuhl, P. Stanat, B. Lütje-Klose, C. Gresch, H. A. Pant & M. Prenzel (Hrsg.), *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 17–39). Wiesbaden: Springer Fachmedien Wiesbaden. [https://doi.org/10.1007/978-3-658-06604-8\\_1](https://doi.org/10.1007/978-3-658-06604-8_1)
- Hellmich, F. & Görel, G. (2014). Erklärungsfaktoren für Einstellungen von Lehrerinnen und Lehrern zum inklusiven Unterricht in der Grundschule. *Zeitschrift für Bildungsforschung, 4*, 227–240. <https://doi.org/10.1007/s35834-014-0102-z>
- Hu, L.-t. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis. Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1–55. <https://doi.org/10.1080/10705519909540118>
- Jackson, D. L., Gillaspy, J. A. & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis. An overview and some recommendations. *Psychological Methods, 14*, 6–23. <https://doi.org/10.1037/a0014694>
- Kemper, C. J., Beierlein, C., Bensch, D., Kovalova, A. & Rammstedt, B. (2012). *Eine Kurzsкала zur Erfassung des Gamma-Faktors sozial erwünschten Antwortverhaltens. Die Kurzsкала Soziale Erwünschtheit-Gamma (KSE-G)* (GESIS – Leibniz-Institut für Sozialwissenschaften, Hrsg.) (GESIS-Working Papers Nr. 25). Mannheim. Verfügbar unter <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-339589>
- Kopp, B. (2009). Inklusive Überzeugung und Selbstwirksamkeit im Umgang mit Heterogenität. Wie denken Studierende des Lehramts für Grundschulen? *Empirische Sonderpädagogik, 1*, 5–25.
- Leyser, Y., Kapperman, G. & Keller, R. (1994). Teacher attitudes toward mainstreaming: A cross-cultural study in six nations. *European Journal of Special Needs Education, 9*, 1–15.
- LimeSurvey 2.65.1 (2016) [Computer software]. Hamburg: LimeSurvey Project Team. Verfügbar unter <http://www.limesurvey.org/>
- Lui, M., Sin, K.-F., Yang, L., Forlin, C. & Ho, F.-C. (2015). Knowledge and perceived social norm predict parents' attitudes towards inclusive education. *International Journal of Inclusive Education, 19*, 1052–1067. <https://doi.org/10.1080/13603116.2015.1037866>
- Lücke, T. (2017) Inclusion ST-IAT. An open (CC-BY) Single Target-Implicit Association Test on Inclusive Education [Computer software]. Dortmund. Verfügbar unter <https://doi.org/10.6084/m9.figshare.4595284>
- Lücke, T. & Grosche, M. (2017a). Implicitly measuring attitudes towards inclusive education. A new attitude test based on single-target implicit associations. *European Journal of Special Needs Education, Advance online article*, 1–10. <https://doi.org/10.1080/08856257.2017.1334432>

- Lüke, T. & Grosche, M. (2017b). *Professions-unabhängige Einstellungsskala zum Inklusiven Schulsystem (PREIS)*. Wuppertal: Bergische Universität Wuppertal. <https://doi.org/10.6084/m9.figshare.2245630>
- Lüke, T. & Grosche, M. (2018). What do I think about inclusive education? It depends on who is asking. Experimental evidence for a social desirability bias in attitudes towards inclusion. *International Journal of Inclusive Education*, 22, 38-53. <https://doi.org/10.1080/13603116.2017.1348548>
- Lüke, T., Przibilla, B., Krämer, P. & Patzelt, M. (2017). *Instrumente zur Messung von Einstellungen zu Inklusion*. [Datenbank]. Wuppertal: Bergische Universität Wuppertal. <https://doi.org/10.6084/m9.figshare.5588584>
- Nilholm, C., Göransson, K. & Patzelt, M., (2016). What is meant by inclusion? An analysis of European and North American journal articles with high impact. *European Journal of Special Needs Education*, 32, 437–451. <https://doi.org/10.1080/08856257.2017.1295638>
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, D. E. Wiley & S. Messick (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49–69). Mahwah, NJ: Erlbaum.
- Paulhus, D. L. (2007). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (Measures of social psychological attitudes, vol. 1, pp. 17–59). San Diego, CA: Academic Press.
- Piezunka, A., Schaffus, T. & Grosche, M. (2017). Vier Definitionen von schulischer Inklusion und ihr konsensueller Kern. Ergebnisse von Experteninterviews mit Inklusionsforschenden. *Unterrichtswissenschaft*, 45, 207-222.
- Przibilla, B., Lauterbach, A., Boshold, F., Linderkamp, F. & Krezmien, M. (2016). Entwicklung und Validierung eines Online-Surveys zur Erhebung von Kompetenzen und Einstellungen von Lehrkräften bezüglich der Inklusion. *Empirische Sonderpädagogik*, 8, 36–53.
- R Core Team. (2016) R: A Language and Environment for Statistical Computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Verfügbar unter <https://www.R-project.org/>
- Revelle, W. (2017) psych: Procedures for Personality and Psychological Research [Computer software]. Evanston, IL: Northwestern University.
- Rosenberg, M. J. & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson & J. W. Brehm (Eds.), *Attitude organization and change. An analysis of consistency among attitude components* (pp. 1–14). New Haven, CT: Yale University Press.
- Rosseel, Y. (2012). lavaan: An R Package for structural equation modeling. *Journal of Statistical Software*, 48, 1–36. Verfügbar unter <http://www.jstatsoft.org/v48/i02/>
- RStudio Team. (2016) RStudio [Computer software]. Boston, MA: RStudio Inc. Verfügbar unter <http://www.rstudio.com/>
- Ruberg, C. & Porsch, R. (2017). Einstellungen von Lehramtsstudierenden und Lehrkräften zur schulischen Inklusion. Ein systematisches Review deutschsprachiger Forschungsarbeiten. *Zeitschrift für Pädagogik*, 63, 393-415.
- Schwab, S. & Feyerer, E. (Hrsg.). (2016) Einstellungsforschung zum inklusiven Unterricht [Themenheft]. *Empirische Sonderpädagogik*, 8. Lengerich: Pabst Science Publishers.
- Schwab, S., Tretter, T. & Gebhardt, M. (2013). Entwicklung und Überprüfung eines fallbasierten Instruments zur Messung der Einstellung zur schulischen Integration. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 83, 20–32. <https://doi.org/10.2378/vhn2014.art02d>
- Schwarz, N. & Bohner, G. (2001). The construction of attitudes. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of so-*

- cial psychology. *Intraindividual Processes* (Handbooks of social psychology, pp. 436–457). Malden, MA: Blackwell Publishers.
- Seifried, S. (2015). *Einstellungen von Lehrkräften zu Inklusion und deren Bedeutung für den schulischen Implementierungsprozess – Entwicklung, Validierung und strukturgeleichungsanalytische Modellierung der Skala EFI-L*. Dissertation, Pädagogische Hochschule Heidelberg, 2015. Heidelberg.
- Seifried, S. & Heyl, V. (2016). Konstruktion und Validierung eines Einstellungsfragebogens zu Inklusion für Lehrkräfte (EFI-L). *Empirische Sonderpädagogik*, 8, 22–35.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's Alpha. *Psychometrika*, 74, 107–120. <https://doi.org/10.1007/s11336-008-9101-0>
- Trumpa, S., Janz, F., Heyl, V. & Seifried, S. (2014). Einstellungen zu Inklusion bei Lehrkräften und Eltern – Eine schulartspezifische Analyse. *Zeitschrift für Bildungsforschung*, 4, 241–256.
- Uziel, L. (2010). Rethinking social desirability scales. From impression management to interpersonally oriented self-control. *Perspectives on psychological science: a journal of the Association for Psychological Science*, 5, 243–262. <https://doi.org/10.1177/1745691610369465>
- Wetzel, E., Böhnke, J. R. & Brown, A. (2016). Response biases. In F. T. L. Leong, D. Bartram, F. M. Cheung, K. F. Geisinger & D. Iliescu (Eds.), *The ITC international handbook of testing and assessment* (pp. 349–363). Oxford, UK: Oxford University Press.

### **Timo Lücke**

Bergische Universität Wuppertal  
 Institut für Bildungsforschung  
 Gaußstr. 20  
 42119 Wuppertal  
 luecke@uni-wuppertal.de

Erstmalig eingereicht: 20.06.2017

Überarbeitung eingereicht: 09.11.2017

Angenommen: 28.11.2017

Wir danken Pia Hansbuer, Christoph Langen, Marco Patzelt, Kristina Pickhardt und Jon Wischmann für ihre Unterstützung bei der Datenerhebung.