

Mechler, Christina; Scheer, David; Heyl, Vera

Psychometrische Qualität von Messinstrumenten zu Sichtweisen auf Inklusion (PREIS, PREIS-E, EFI-L, FEDI)

Empirische Sonderpädagogik 16 (2024) 2, S. 137-153



Quellenangabe/ Reference:

Mechler, Christina; Scheer, David; Heyl, Vera: Psychometrische Qualität von Messinstrumenten zu Sichtweisen auf Inklusion (PREIS, PREIS-E, EFI-L, FEDI) - In: Empirische Sonderpädagogik 16 (2024) 2, S. 137-153 - URN: urn:nbn:de:0111-pedocs-322235 - DOI: 10.25656/01:32223; 10.2440/003-0025

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-322235>

<https://doi.org/10.25656/01:32223>

in Kooperation mit / in cooperation with:

Pabst Science Publishers <https://www.psychologie-aktuell.com/journale/empirische-sonderpaedagogik.html>

Nutzungsbedingungen

Dieses Dokument steht unter folgender Creative Commons-Lizenz: <http://creativecommons.org/licenses/by-nc/4.0/deed.de> - Sie dürfen das Werk bzw. den Inhalt vervielfältigen, verbreiten und öffentlich zugänglich machen sowie Abwandlungen und Bearbeitungen des Werkes bzw. Inhaltes anfertigen, solange Sie den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen und das Werk bzw. den Inhalt nicht für kommerzielle Zwecke verwenden.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

This document is published under following Creative Commons-License: <http://creativecommons.org/licenses/by-nc/4.0/deed.en> - You may copy, distribute and render this document accessible, make adaptations of this work or its contents accessible to the public as long as you attribute the work in the manner specified by the author or licensor. You are not allowed to make commercial use of the work, provided that the work or its contents are not used for commercial purposes.

By using this particular document, you accept the above-stated conditions of use.



Kontakt / Contact:

peDOCS

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation
Informationszentrum (IZ) Bildung

E-Mail: pedocs@dipf.de

Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

Empirische Sonderpädagogik, 2024.16:137-153
DOI <https://doi.org/10.2440/003-0025>
ISSN 1869-4845 (Print) · ISSN 1869-4934 (ebook)

Psychometrische Qualität von Messinstrumenten zu Sichtweisen auf Inklusion (PREIS, PREIS-E, EFI-L, FEDI)

Christina Mechler¹, David Scheer² & Vera Heyl¹

¹ Annelie-Wellensiek-Zentrum für Inklusive Bildung, Pädagogische Hochschule Heidelberg

² Pädagogische Hochschule Ludwigsburg

³ Eugen-Neter-Schule Mannheim, Sonderpädagogisches Bildungs- und Beratungszentrum mit dem Förderschwerpunkt Geistige Entwicklung

Zusammenfassung

Sichtweisen auf Inklusion werden im Kontext der Professionalisierung von (angehenden) Lehrkräften seit einigen Jahren intensiv beforscht. Dabei entstand eine Vielzahl an Instrumenten, um Einstellungen und subjektive Sichtweisen zu untersuchen, deren psychometrische Qualität in sehr unterschiedlicher Weise geprüft wurde. Die Sicherstellung dieser in Form von Replikationen findet bislang wenig Eingang in den sonder- und inklusionspädagogischen Forschungsdiskurs. Im Fokus des Beitrags steht die Replikation der Gütekriterien der Instrumente PREIS, PREIS-E, EFI-L und FEDI mit Lehramtsstudierenden aller Schularten ($N = 450$). Die interne Konsistenz sowie Retest-Reliabilität aller untersuchten Skalen ist akzeptabel bis gut mit Ausnahme des FEDI. Die Ergebnisse konfirmatorischer Faktorenanalysen bestätigen die postulierten Faktorenstrukturen, wenngleich bei EFI-L, PREIS und FEDI einzelne Modellfit-Indizes Schwächen aufweisen. Signifikante Korrelationen zwischen den untersuchten Instrumenten liefern weitere Hinweise auf deren Konstruktvalidität. Darüber hinaus liefern die Ergebnisse Hinweise auf Messinvarianz zwischen unterschiedlichen Lehramtsstudiengängen.

Schlagwörter: Einstellungen zu Inklusion, Einstellungsmessung, konfirmatorische Faktorenanalyse, Reliabilität, Replikationsstudie

Psychometric Properties of Instruments Measuring Views on Inclusion (PREIS, PREIS-E, EFI-L, FEDI)

Summary

In the field of teacher education research, attitudes towards inclusion (ATI) have been investigated for some years. Many different ATI measures resulted from this research as well as a variety of approaches to examine the psychometric quality of these measures. Yet, replication studies to assure findings have been neglected in special needs and inclusive educational research. This article aims to investigate the psychometric properties of the (German) measures PREIS, PREIS-E, EFI-L and FEDI based on a sample of pre-service teachers ($N = 450$). Internal consistency and retest reliability is acceptable up to good for all

measures, except for FEDI. Results from confirmatory factor analyses support the postulated factor structures, although for EFI-L, PREIS and FEDI, some of the model-fit-indices indicate shortcomings. Significant correlations between all measures indicate construct validity. Furthermore, the results indicate measurement invariance between teacher education programs.

Keywords: Attitudes towards Inclusion, Attitude Measurement, Confirmatory Factor Analysis, Reliability, Replication Study

Die Untersuchung von Einstellungen zu Inklusion als Aspekt der inklusiven Lehrkräftekompetenz stellt seit einigen Jahren einen Forschungsschwerpunkt der Sonder- und Inklusionspädagogik dar (Avramidis & Norwich, 2002; de Boer et al., 2011). Insgesamt liegt eine Vielzahl an Studien zu Einstellungen zu Inklusion vor. Eine Schwierigkeit bei der Interpretation der Ergebnisse liegt allerdings darin, dass die Begriffe ‚Einstellung‘, ‚Überzeugung‘ oder ‚belief‘ in bisherigen Forschungsarbeiten nicht immer eindeutig voneinander abgegrenzt oder uneinheitlich definiert werden (Ruberg & Porsch, 2017). *Einstellungen* werden im vorliegenden Beitrag als wertende Urteile über ein Einstellungsobjekt verstanden, die sich hinsichtlich ihrer Valenz und Stärke unterscheiden können und kognitive, affektive und verhaltensbezogene Informationen enthalten (Jonas et al., 2014). Einem solchen Verständnis folgen auch mehrere Messinstrumente zur Erfassung von Einstellungen zu Inklusion (z. B. EFI-L, Seifried & Heyl, 2016; PREIS, Lüke & Grosche, 2018). Verbunden mit der Forschung zu Einstellungen ist auch die Frage nach deren Bedeutung für Verhalten. Entsprechend der *Theorie des geplanten Verhaltens* entsteht Verhalten als Resultat einer Verhaltensintention, die wiederum durch drei Elemente vorhergesagt werden kann (Ajzen, 1991):

1. Einstellungen zu einem bestimmten Verhalten, bspw. zu inklusivem Unterricht,
2. die subjektiv wahrgenommene soziale Norm (z.B. inwiefern subjektiv bedeutsame Personen Inklusion unterstützen)
3. sowie die wahrgenommene Verhaltenskontrolle (also, ob eine Person sich unter

Berücksichtigung vorheriger Erfahrungen und angenommener Hindernisse zutraut, bspw. inklusiv zu unterrichten). Die wahrgenommene Verhaltenskontrolle ist kompatibel mit dem Konzept der Selbstwirksamkeitserwartung nach Bandura (Ajzen, 1991).

Das Zusammenspiel dieser drei Elemente bedingt die Verhaltensintention in der Weise, dass diese umso stärker ist, je positiver die Einstellungen und die subjektiv wahrgenommene soziale Norm und je höher die subjektiv wahrgenommene Verhaltenskontrolle sind (Ajzen, 1991). Dementsprechend können die drei unterschiedlichen Konstrukte im Kontext der Inklusionsforschung Auskunft über die Verhaltensintention zu Inklusion geben. Insbesondere die Zusammenhänge von Einstellungen und Selbstwirksamkeitserwartungen im Hinblick auf Inklusion wurden vielfach untersucht, wobei in einer Metaanalyse bisheriger Forschungserkenntnisse ein mittelstarker Zusammenhang der Konstrukte bestätigt werden konnte (Yada et al. 2022). Weiterhin wurden in mehreren Studien signifikante Korrelationen zwischen subjektiv wahrgenommener sozialer Norm und Einstellungen festgestellt ($r = .17-.54$), während der Zusammenhang zwischen subjektiv wahrgenommener sozialer Norm und wahrgenommener Verhaltenskontrolle gemischte Befunde aufweist ($r = .09-.50$; MacFarlane & Woolfson, 2013; Wilson et al., 2016; Hellmich et al., 2019). Bislang liegen kaum Messinstrumente vor, die alle drei Elemente der Theorie des geplanten Verhaltens bezogen auf Inklusion erfassen. Eine Ausnahme

bildet die PREIS-E (Lüke & Grosche, 2019), die als Erweiterung des Einstellungsfragebogens PREIS (Lüke & Grosche, 2018) die subjektiv wahrgenommene soziale Norm und Verhaltenskontrolle erfasst.

Subjektive Definitionen von Inklusion hingegen können im Kontext subjektiver Theorien von Lehrkräften betrachtet werden (Mandl & Huber, 1983). Mandl und Huber (1983) definieren diese als Kognitionen von Lehrkräften, die deren subjektive Sichtweise des Erlebens und Handelns widerspiegeln. Sie sind in ihrer Struktur mit einer wissenschaftlichen Theorie vergleichbar und beinhalten individuelle Überzeugungen und Interpretationen, Schlussfolgerungen, Handlungsmöglichkeiten und –gewohnheiten, die auch Alltagswissen und –konzepte miteinschließen (Mandl & Huber, 1983). Subjektive Definitionen zu Inklusion umfassen demzufolge, was eine Person unter Inklusion versteht, während Einstellungen die Bewertung von Inklusion enthalten. Im Kontext der Einstellungsforschung zu Inklusion können subjektive Definitionen auf Einstellungen zu Inklusion einen Einfluss haben, da Einstellungen entsprechend dem zugrundeliegenden Verständnis von Inklusion unterschiedlich ausfallen können (Krischler et al., 2019). Bislang gibt es wenige Forschungserkenntnisse zu diesem Konstrukt, allerdings konnten in einer ersten explorativen Studie bei Lehramtsstudierenden sehr uneinheitliche subjektive Definitionen von Inklusion mit teils widersprüchlichen Elementen herausgearbeitet werden (Przibilla et al., 2018). Auch entsprechende Messinstrumente liegen bislang nur vereinzelt vor (z. B. FEDI, Scheer et al., 2023).

Messinstrumente zu Sichtweisen auf Inklusion

Im Rahmen der Vielzahl an Studien in den letzten Jahren wurde eine Vielfalt an Messinstrumenten (weiter-)entwickelt, um Sichtweisen auf Inklusion zu erfassen. Beim Vergleich der vorliegenden Messinstrumente lässt sich eine große Heterogenität bezogen

auf folgende Aspekte feststellen: Erstens lassen sich unterschiedliche theoretische Bezüge erkennen, da den Instrumenten verschiedene Einstellungskonstrukte (Ruberg & Porsch, 2017) oder Inklusionsverständnisse (Grosche & Lüke, 2020) zugrunde liegen. Dies zeigt sich z.B. darin, dass einige Einstellungsinstrumente von einem Dreikomponentenmodell ausgehen (Eagly & Chaiken, 2010), wohingegen andere diese als ein inhaltliches Evaluationsmuster betrachten, bei dem keine einzelnen Komponenten abgrenzbar sind (Lüke & Grosche, 2018). Zweitens unterscheiden sich die inhaltlichen Schwerpunkte der Instrumente, z.B. indem sie sich auf ein inklusives Schulsystem (z.B. PREIS; Lüke & Grosche, 2018), konkret auf den inklusiven Unterricht (z.B. EFI-L; Seifried & Heyl, 2016) oder nicht auf Einstellungen im engeren Sinne, sondern eher auf subjektive Definitionen zu Inklusion (z.B. FEDI; Scheer et al., 2023) beziehen. Und drittens sind teils unterschiedliche Personengruppen adressiert: Während ein Schwerpunkt in der Einstellungsmessung von (angehenden) Lehrkräften (z.B. EFI-L; Seifried & Heyl, 2016) liegt, gibt es auch Instrumente, die weitere Personengruppen umfassen oder professionsunabhängig (z.B. PREIS; Lüke & Grosche, 2018) konzipiert sind.

In der vorliegenden Studie wurden Instrumente gewählt, die unterschiedliche Facetten von Sichtweisen erfassen, um die Gütekriterien einer möglichst großen Bandbreite von Instrumenten zu überprüfen: Zur Untersuchung der Einstellungen zu Inklusion wurden der *Einstellungsfragebogen zu Inklusion für Lehrkräfte* (Seifried & Heyl 2015) und die *Professionsunabhängige Einstellungsskala zum inklusiven Schulsystem* (PREIS, Lüke & Grosche 2018) eingesetzt. Beide Instrumente wurden seit ihrer Entwicklung im deutschsprachigen Raum in vielen Studien eingesetzt, die Fragestellungen zu Einstellungen zu Inklusion beinhalten, und sollten deshalb studienübergreifend eine stabile psychometrische Qualität aufweisen, um Forschungsergebnisse angemessen interpretieren zu können. Die *Erweiterung der PREIS*

(PREIS-E, Lüke & Grosche 2019) wurde zur Erfassung der subjektiv wahrgenommenen Verhaltenskontrolle und sozialen Norm als weitere Bestandteile der Verhaltensintention zu Inklusion verwendet. In Verbindung mit der PREIS gibt sie Aufschluss über alle drei Bereiche der Theorie des geplanten Verhaltens, jedoch sind uns keine Studien bekannt, in denen das Instrument eingesetzt bzw. dessen Testgüte untersucht wurde. Die subjektiven Definitionen von Inklusion wurden mit dem *Fragebogen zur Erfassung subjektiver Definitionen von Inklusion* (FEDI; Scheer et al. 2023) untersucht, welches das erste Instrument zu diesem Konstrukt darstellt und zur Sicherstellung der Testgüte weiterer Untersuchung bedarf.

Evaluation der Testgüte von EFI-L, PREIS, PREIS-E und FEDI

Zum *EFI-L* (Seifried & Heyl, 2016) liegt eine Validierungs- sowie eine erste Replikationsstudie mit unbeteiligten Forschenden vor, bei der die Reliabilität ($\alpha = .81-.90$) und die Faktorenstruktur des Instruments ($\chi^2/df = 3.23-4.6$, $TLI = .93$, $CFI = .94$, $RMSEA = .06-.07$) bestätigt werden konnten (Gorges et al., 2020; Seifried & Heyl, 2016). Die Itemladungen lagen zwischen .56 und .84, die Subskalen korrelierten zwischen .45 und .78. Gorges et al. (2020) konnten überdies hohe Korrelationen mit zwei anderen Instrumenten zu Einstellungen zu Inklusion ($r = .55-.77$) sowie moderate bis hohe Korrelationen mit der Selbstwirksamkeit bezüglich inklusiven Unterrichts ($r = .29-.61$) berechnen. Zusätzlich geben fehlende bis schwache Zusammenhänge Hinweise auf diskriminante Validität von Einstellungen zu Inklusion und Persönlichkeitsmerkmalen (Big Five). Konfigurale und schwache, nicht jedoch starke Messinvarianz konnte im Vergleich von Lehramtsstudierenden und Lehrkräften bestätigt werden (Gorges et al., 2020).

Die psychometrische Qualität der *PREIS* (Lüke & Grosche, 2018) wurde in einer Validierungs- und einer Replikationsstudie geprüft, die die Reliabilität (Retest-Relia-

bität: $r_{tt} = .94-.96$; interne Konsistenz: $\alpha = .93$) und das einfaktorielles Modell ($\chi^2/df = 1.76-2.51$, $CFI = .96-.97$, $TLI = .95-.97$, $RMSEA = .05-.06$) bestätigen konnten (Lüke & Grosche, 2018; Schulze et al., 2019). Die Itemladungen lagen zwischen .39 und .81. Die *PREIS* korreliert moderat mit Selbstwirksamkeit bezüglich inklusiven Unterrichts ($r = .28$). In beiden Studien waren teils dieselben Forschenden beteiligt.

Zur Erweiterung der *PREIS* (*PREIS-E*, Lüke & Grosche, 2019) liegt bislang keine Studie vor, die die Güte des Instruments evaluiert hat.

Zum *FEDI* (Scheer et al., 2023) liegt eine Validierungsstudie vor, die die Reliabilität ($\alpha = .70-.79$) und die Faktorenstruktur ($\chi^2/df = 2.32$, $CFI = .97$, $TLI = .96$, $RMSEA = .07$) des Instruments analysiert (Scheer et al., 2023). Die Itemladungen variierten zwischen .41 und .86, die Subskalen zeigten keine bis moderate Zusammenhänge ($r = .09-.36$). Faktorenstruktur, Reliabilität sowie Korrelation mit Einstellungen zu Inklusion des *FEDI* konnten auch in einer für den frühpädagogischen Bereich adaptierten Version bestätigt werden, die von den gleichen Forschenden entwickelt wurde (Scheer & Melzer, in Vorb.). Darüber hinaus konnten Scheer et al. (2023) geringe bis moderate Korrelationen der subjektiven Definitionen mit Einstellungen zu Inklusion ($r = .20-.44$) und geringe Zusammenhänge mit der Selbstwirksamkeit bezüglich Inklusion berechnen ($r = .14-.24$), was sich auch in der adaptierten Fassung bestätigte (Scheer & Melzer, in Vorb.). Zudem konnte starke Messinvarianz bestätigt werden zwischen Studierenden der Sonderpädagogik und anderer Lehramtsstudiengänge (Primar- und Sekundarstufe; Scheer et al., 2023).

Zusammenfassend lässt sich sagen, dass zur Evaluation der Güte der Instrumente *EFI-L*, *PREIS* und *FEDI* in Validierungsstudien sowohl Reliabilitäts- und konfirmatorische Faktorenanalysen als auch Korrelationsanalysen durchgeführt wurden, während diese Analysen bei *EFI-L* und *PREIS* zusätzlich in Replikationsstudien vollzogen wurden. Darüber hinaus wurde die Messinvarianz von

EFI-L (zwischen Lehramtsstudierenden und Lehrkräften) und FEDI (zwischen Sonderpädagogik- und anderen Lehramtsstudierenden) geprüft. Zum PREIS-E liegt bislang keine Studie vor, die sich mit der Evaluation der Güte befasst.

Replikationsstudien in der Erziehungswissenschaft

Als Replikationen werden Studien bezeichnet, die Ergebnisse oder Hypothesen in Form eines Wiederholungsprozesses bestätigen bzw. verwerfen und somit eine Reliabilitätsprüfung vorheriger Forschungsergebnisse ermöglichen (Plucker & Makel, 2021; Schmidt, 2009). Im Zuge der Replikationskrise wurden in verschiedenen wissenschaftlichen Disziplinen Fragestellungen zur Replikation von Forschungsergebnissen untersucht. Im Kontext der Erziehungswissenschaften kamen Makel und Plucker (2014) in einer Metaanalyse zum Ergebnis, dass lediglich 0,13% der Forschungsergebnisse, die in den 100 top-gerankten Zeitschriften für Erziehungswissenschaften veröffentlicht wurden, auf Replikationsstudien basieren (Makel & Plucker, 2014). Insgesamt konnten die Ergebnisse der Originalarbeiten bei 67% der Studien repliziert werden. Dabei scheint die (Nicht-) Beteiligung von Forschenden aus Original- und Replikationsstudie einen Einfluss auf die Replikationsrate und Replizierbarkeit zu haben: Knapp die Hälfte der eingeschlossenen Studien (48,2%) wurde vom selben Forschungsteam durchgeführt. Bei diesen Studien konnte eine Replikationsrate von 88,7% erzielt werden, während bei der Beteiligung von mindestens einem Forschenden in beiden Studien eine Replikationsrate von 70,6% erreicht wurde. Bei fehlender Überlappung der beteiligten Personen waren 54% der Replikationen gelungen (Makel & Plucker, 2014).

Ein Desiderat stellt zusammenfassend im Kontext vieler sozialwissenschaftlicher Disziplinen die Veröffentlichung von Studien dar, die veröffentlichte Forschungsergebnisse erneut prüfen, sowie von Replikationen, welche unabhängig von Forschenden

aus den Originalstudien durchgeführt werden (Plucker & Makel, 2021). Die Relevanz zeigt sich im Fall von Messinstrumenten in besonderer Weise, da diese häufig in verschiedensten Studien Anwendung finden und deshalb valide und reliabel die postulierten Konstrukte erfassen sollten. Somit hängt die Gültigkeit der berichteten Ergebnisse mit der Gültigkeit der verwendeten Instrumente zusammen.

Forschungsfrage

Das Desiderat fehlender Replikationen von bereits veröffentlichten Messinstrumenten im Kontext inklusiver Bildung greift der vorliegende Artikel auf, indem eine direkte Replikation der Untersuchung von Instrumenten zur Einstellungsmessung durchgeführt wird: Alle Instrumente adressieren inhaltlich Sichtweisen auf Inklusion, indem die Einstellungen zu Inklusion bezüglich eines inklusiven Schulsystems und inklusiven Unterrichts sowie die Verhaltensintention zu Inklusion und subjektive Definitionen von Inklusion erfasst werden.

Die Untersuchungsschwerpunkte liegen entsprechend den Methoden der Originalstudien in der Replizierbarkeit der Faktorenstrukturen durch konfirmatorische Faktorenanalysen sowie der Prüfung von Reliabilität (interne Konsistenz und Retest Reliabilität) sowie Validität der Instrumente (konvergente und diskriminante Validität über Korrelationsanalysen). Darüber hinaus wurde die Messinvarianz bezogen auf den Studiengang (Primarstufe, Sekundarstufe I/II, Sonderpädagogik) geprüft. Teils handelt es sich um eine erste Replikation von Originalarbeiten (PREIS-E, FEDI), teils um die Prüfung der psychometrischen Qualität als „ongoing process“ (EFI-L, PREIS), da bereits Replikationen der Originalarbeiten vorliegen. Dementsprechend liegt diesem Artikel die Forschungsfrage zugrunde:

Inwiefern lässt sich die postulierte psychometrische Qualität der Instrumente zu Einstellungen zu Inklusion (EFI-L, Seifried

& Heyl, 2016; PREIS, Lüke & Grosche, 2018; PREIS-E, Lüke & Grosche, 2019;) und zu subjektiven Definitionen von Inklusion (FEDI, Egner et al., 2019; Scheer et al., 2023) bestätigen?

Methoden

Stichprobe und Procedere

Die Studie wurde im Sommersemester 2021 bei Lehramtsstudierenden aus fünf Hochschulen durchgeführt ($N = 450$). Die Daten sind Teil einer Interventionsstudie, wobei hauptsächlich die Daten des ersten Messzeitpunktes für die vorliegende Studie berücksichtigt wurden. Alle Teilnehmenden besuchten in diesem Semester eine Hochschullehrveranstaltung zu den Themen Diversität, Heterogenität oder Inklusion, in deren Rahmen die Erhebungen durchgeführt wurden. Für einen Teil der Analysen wurden auch die Datensätze des zweiten Messzeitpunktes der Interventionsstudie verwendet, in diesem Fall jedoch ausschließlich der Vergleichs-, nicht der Interventionsgruppe ($n = 68$). Die zweite Erhebung fand mit einem zeitlichen Abstand von ca. drei Monaten statt. Aufgrund der Corona-Pandemie fanden alle Erhebungen online statt. Die Erhebung wurde auf der Plattform Limesurvey (Limesurvey GmbH, 2023) durchgeführt. Ein positives Votum der Ethikkommission liegt vor. Darüber hinaus wurden personenbezogene Daten der Teilnehmenden (Geschlecht, Alter, Studiengang, Studienphase) erfasst. Die Zusammensetzung der Stichprobe lässt sich Tabelle 1 entnehmen.

Instrumente

Einstellungsfragebogen zu Inklusion für Lehrkräfte

Der Einstellungsfragebogen zu Inklusion für Lehrkräfte (EFI-L; Gorges et al., 2020; Seifried & Heyl, 2016) basiert auf einem dreifaktoriellen Einstellungsmodell (Eagly &

Chaiken 2010) und setzt sich mit insgesamt 15 Items aus den drei Faktoren „Fachliche Förderung im inklusiven Unterricht“ (z.B. „Kinder mit besonderen Bedürfnissen würden in einer inklusiven Schulklasse letztendlich besser gefördert“), „Persönliche Bereitschaft zu inklusivem Unterricht“ (z.B. „Der Unterricht in einer inklusiven Klasse macht mir Angst“) sowie „Soziale Inklusion im Unterricht“ (z.B. „Kinder mit besonderen Bedürfnissen würden in einer inklusiven Schulklasse von den anderen Kindern gut behandelt werden“) zusammen. Die Skala zeigt sechs ausformulierte Stufen mit den Endpunkten *stimme gar nicht zu* bis *stimme voll und ganz zu*.

Professionsunabhängige Einstellungsskala zum Inklusiven Schulsystem

Ein erstes Instrument mit dem Schwerpunkt Einstellungen zum inklusiven Schulsystem bei verschiedenen Professionen ist die Professionsunabhängige Einstellungsskala zum Inklusiven Schulsystem (PREIS; Lüke & Grosche, 2018; Schulze et al., 2019). Die einfaktorische Skala enthält 14 Items (z.B. „Der Gedanke an ein inklusives Schulsystem stimmt mich fröhlich“) bei einem fünfstufigen Antwortformat mit den formulierten Endpunkten *stimme nicht zu* und *stimme zu*.

In Ergänzung zur PREIS wurde die *Erweiterung PREIS-E* (Lüke & Grosche, 2019) eingesetzt, die die Konstrukte „wahrgenommene soziale Norm“ (z.B. „Menschen, die mir wichtig sind, halten ein inklusives Schulsystem für sinnvoll“) und die „wahrgenommene Verhaltenskontrolle“ (z.B. „Ich habe keinen Einfluss auf die Gestaltung des Schulsystems“) im ebenfalls fünfstufigen Antwortformat mit insgesamt acht Items erfasst. Diese Konstrukte knüpfen an die Theorie des geplanten Verhaltens (Ajzen, 1991) an.

Tabelle 1. Zusammensetzung der Stichprobe

Variable	Ausprägung	MZP 1 N	in %	MZP 2 N	in %
N		450		68	
Geschlecht	Weiblich	392	87.1	61	89.7
	Männlich	55	12.2	6	8.8
	fehlend	3	0.6	1	1.5
Studiengang	Lehramt Primarstufe	214	47.6	47	69.1
	Lehramt Sek I + Sek II	100	22.3	9	13.2
	Lehramt Sonderpädagogik	135	30.0	12	17.6
	fehlend	1	0.2	-	-
Studienphase	Bachelor	66	59.1	34	50
	Master/Staatsexamen	181	40.2	33	48.5
	fehlend	3	0.7	1	1.5
Alter [M [SD]]		23.95 [4.15]		24.09 [4.57]	

Fragebogen zur Erfassung subjektiver Definitionen von Inklusion

Die subjektiven Definitionen von Inklusion wurden mit dem Fragebogen zur Erfassung subjektiver Definitionen von Inklusion erhoben (FEDI; Egner et al., 2019; Scheer et al., 2023). Dieser besteht aus insgesamt 15 Items mit den drei Subskalen „Schulsystemische Perspektive“ (z.B. „In einem inklusiven Schulsystem gibt es keine speziellen Förderschulen“), „Menschenrechtsbasierte ethische Perspektive“ (z.B. „Inklusion heißt, Menschen dazu zu ermächtigen, für ihre Rechte einzutreten“) und „Förderungsorientierte Perspektive“ (z.B. „Bei Inklusion steht die Lernentwicklung der Schülerinnen und Schüler im Mittelpunkt“) in einem fünfstufigen Antwortformat mit formulierten Stufen und den Endpunkten *trifft gar nicht zu* bis *trifft völlig zu*.

Auswertung

Zur Auswertung der Daten wurden die Programme IBM SPSS (IBM Corp., 2022b), IBM SPSS AMOS (IBM Corp., 2022a) und R (R Core Team, 2023) mit dem Paket lavaan (Rosseel, 2023) verwendet.

Die Reliabilität wurde erstens über die interne Konsistenz ermittelt. Zweitens wurde die Retest-Reliabilität berechnet, wobei nur ein Teil der Datensätze verwendet wurde ($n = 68$, s. Tab. 1). Für die Reliabilitätsanalysen wurden unvollständige Fälle paarweise ausgeschlossen.

Die jeweils postulierten Faktorenstrukturen, die im Abschnitt Instrumente und den Abbildungen 1–4 nachvollzogen werden können, wurden über konfirmatorische Faktorenanalysen geprüft. Dafür wurden zunächst negativ formulierte Items invertiert. Für die Schätzung der Parameter wurde die *Full Information Maximum Likelihood-Methode* angewendet. Dabei sollten die Itemladungen bei einer Ladungsstärke $\lambda > .5$ liegen (Backhaus et al., 2015), während die Ergebnisse der Chi-Quadrat-Teststatistik $\chi^2/df < 3$, der RMSEA $< .08$ und die inkrementellen Fitmaße (*CFI*, *TLI*, *NFI*) $> .9$ sein sollten (Moosbrugger & Kelava, 2020).

Korrelationsanalysen wurden zur Berechnung der Interkorrelationen der untersuchten Instrumente durchgeführt, um weitere Hinweise auf Konstruktvalidität zu erlangen. Zur Berechnung der Mittelwerte wurden diejenigen Fälle eingeschlossen, die

mindestens 80% der Items einer Skala beantwortet hatten.

Die Prüfung der Messinvarianz nach Lehramtsstudiengang erfolgte mittels Multi-Gruppen-CFA in der Schrittfolge konfigurale (Faktorenstruktur über Gruppen identisch), schwache (zusätzlich identische Ladungen), starke (zusätzlich identische Intercepts) und strikte (zusätzlich identische Residuen) Messinvarianz. Messinvarianz wird jeweils angenommen, wenn der Modellfit sich im Vergleich zum vorherigen Schritt nicht signifikant verschlechtert. Da der absolute χ^2 -Test zu sensitiv auf die Stichprobengröße reagiert, wird als Cut-Off-Wert meist eine Verschlechterung von weniger als 0,015 im *RMSEA* bzw. im *CFI* weniger 0,01 (Chen, 2007; Cheung & Rensvold, 2002) bis 0,02 (Rutkowski & Svetina, 2014; Schwab & Helm, 2015) angenommen.

Ergebnisse

Prüfung der Reliabilität

Für alle Skalen bzw. Subskalen wurde die interne Konsistenz berechnet, die in allen Fällen mindestens akzeptabel bis sehr gut ist ($\alpha = .70-.91$, s. Tab. 2), mit Ausnahme der Subskala „schulsystemische Perspektive“ des FEDI ($\alpha = .59$). Auch wurde anhand einer kleineren Stichprobe ($n = 68$) die Retest-Reliabilität berechnet. Während bei PREIS ($r_{tt} = .82$) und bei EFI-L ($r_{tt} = .79$) eine hinreichend hohe Retest-Reliabilität vorliegt, ist sie bei PREIS-E ($r_{tt} = .62-.67$) und FEDI ungenügend ($r_{tt} = .55-.57$; Moosbrugger & Kelava, 2020).

Replikation der Faktorenstruktur

Insgesamt lässt sich über alle Instrumente hinweg feststellen, dass das jeweilige Gesamtmodell plausibel ist, d.h. alle Indikatoren weisen positive, statistisch signifikant von Null abweichende Ladungen auf. Bezüglich der Ladungsstärke genügen die Indikatoren der Skalen PREIS-E und EFI-L

der Ladungsstärke $\lambda > .5$ (Backhaus et al., 2015), während einige Indikatoren der übrigen Skalen schwächere Ergebnisse aufweisen: In der PREIS konnte ein auffälliges Item identifiziert werden (Item 9: „Ich meine, dass ein inklusives Schulsystem mit dem Leistungsprinzip unserer Gesellschaft nicht vereinbar ist.“, $\lambda = .29$). Beim FEDI laden sieben der 15 Items mit überwiegend knapp unter .5 (Items 1, 2, 4, 8, 10, 13, 15).

Auf Faktorebene weisen die beiden Faktoren der PREIS-E keinen signifikanten Zusammenhang auf ($r = .15$). Beim EFI-L ergeben sich zwischen den drei Faktoren mittelstarke Korrelationen ($r = .37-.58$). Zwischen den Faktoren von FEDI zeigen sich keine bzw. moderate Korrelationen ($r = .03-.40$).

Alle Ergebnisse lassen sich über die Pfadmodelle nachvollziehen (Abb. 1–4).

Die vier Modelle wurden anhand verschiedener Fit Indizes beurteilt (s. Tab. 3). Während die *RMSEA*-Werte aller untersuchten Konstrukte auf einen akzeptablen Modellfit hinweisen (*RMSEA* $< .08$), zeigen EFI-L und PREIS leichte Einbußen in der Chi-Quadrat-Teststatistik ($\chi^2/df > 3$; Moosbrugger & Kelava, 2020). Des Weiteren weisen die inkrementellen Fitmaße (*CFI*, *TLI*, *NFI*) mit Werten $> .9$ bei PREIS, PREIS-E und EFI-L auf einen guten Modellfit hin, während sie bei FEDI unterhalb dieses Cut-offs liegen (Moosbrugger & Kelava, 2020).

Untersuchung der Konstruktvalidität

Die Korrelation zwischen den beiden Instrumenten zu Einstellungen zu Inklusion PREIS und EFI-L ist stark ($r = .74$). Der lineare Zusammenhang der wahrgenommenen Verhaltenskontrolle (Subskala wVK der PREIS-E) mit den Einstellungen bezüglich Inklusion ist moderat (PREIS: $r = .28$; EFI-L: $r = .31$), während die subjektiv wahrgenommene soziale Norm (Subskala wSN der PREIS-E) hoch mit den Instrumenten zu Einstellungen zu Inklusion korreliert (PREIS: $r = .62$; EFI-L: $r = .54$). Zwischen den subjektiven Definitionen zu Inklusion (FEDI) und

Tabelle 2. Reliabilität, deskriptive Statistik, Korrelationen der Konstrukte

Instrument	PREIS	PREIS-E		EFI-L			FEDI			
Subskala		SN	VK	gesamt	FF	PB	SI	SP	MEP	FP
Itemzahl	14	4	4	15	6	5	4	5	5	5
α	.91	.82	.76	.88	.83	.85	.86	.59	.64	.81
r_{tt}	.82**	.67**	.62**	.79**	.75**	.75**	.82**	.57**	.55**	.56**
M [SD]	2.85 [.71]	2.40 [.77]	2.45 [.82]	3.06 [.73]	2.94 [.87]	2.76 [1.07]	3.56 [.83]	2.67 [.66]	3.12 [.57]	2.17 [.72]
PREIS	1	.62**	.28**	.74**	.74**	.51**	.44**	.29**	.30**	.02
PREIS-SN		1	.10*	.54**	.52**	.37**	.34**	.10*	.14**	.07
PREIS-VK			1	.31**	.25**	.28**	.17**	.05	.06	-.08*
EFI-L				1				.26**	.19**	-.01
FEDI-SP					.31**	.08	.21**	1	.25**	.01
FEDI-MEP					.22**	.11*	.13**		1	.14**
FEDI-FP					.11*	-.10*	-.05			1

SN = wahrgenommene soziale Norm, VK = wahrgenommene Verhaltenskontrolle, FF = fachliche Förderung, PB = Persönliche Bereitschaft, SI = soziale Integration, SP = schulsystemische Perspektive, MEP = Menschenrechtsbasiert-ethische Perspektive, FP = Förderungsorientierte Perspektive

Abbildung 1. Pfadmodell EFI-L

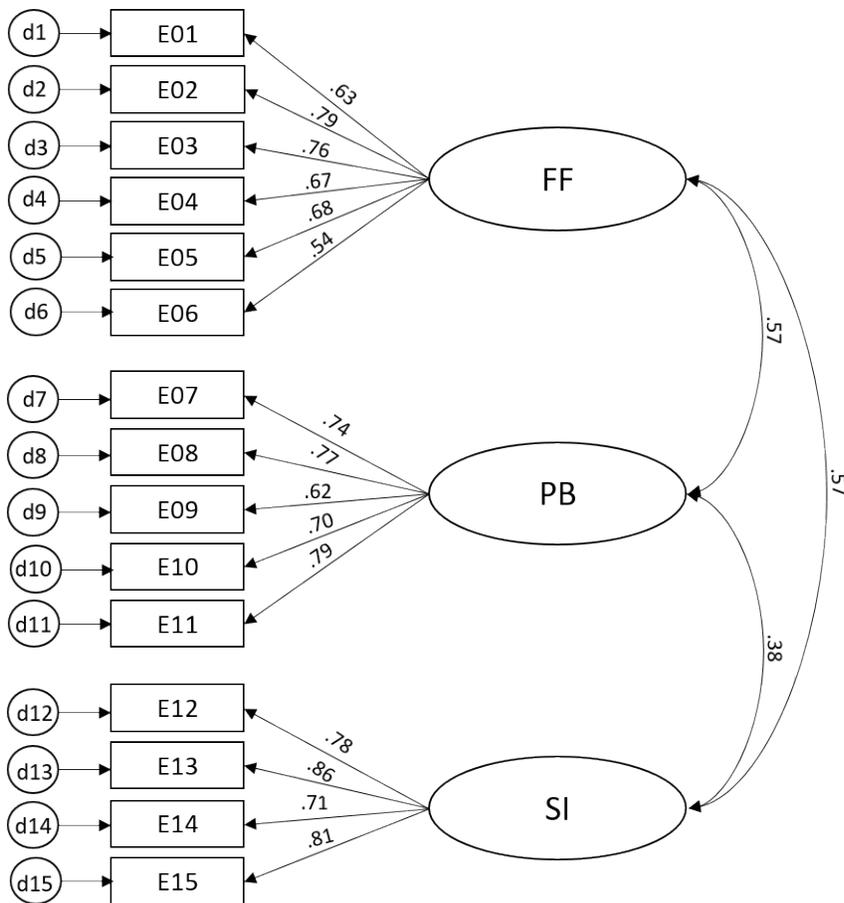


Abbildung 2. Pfadmodell PREIS

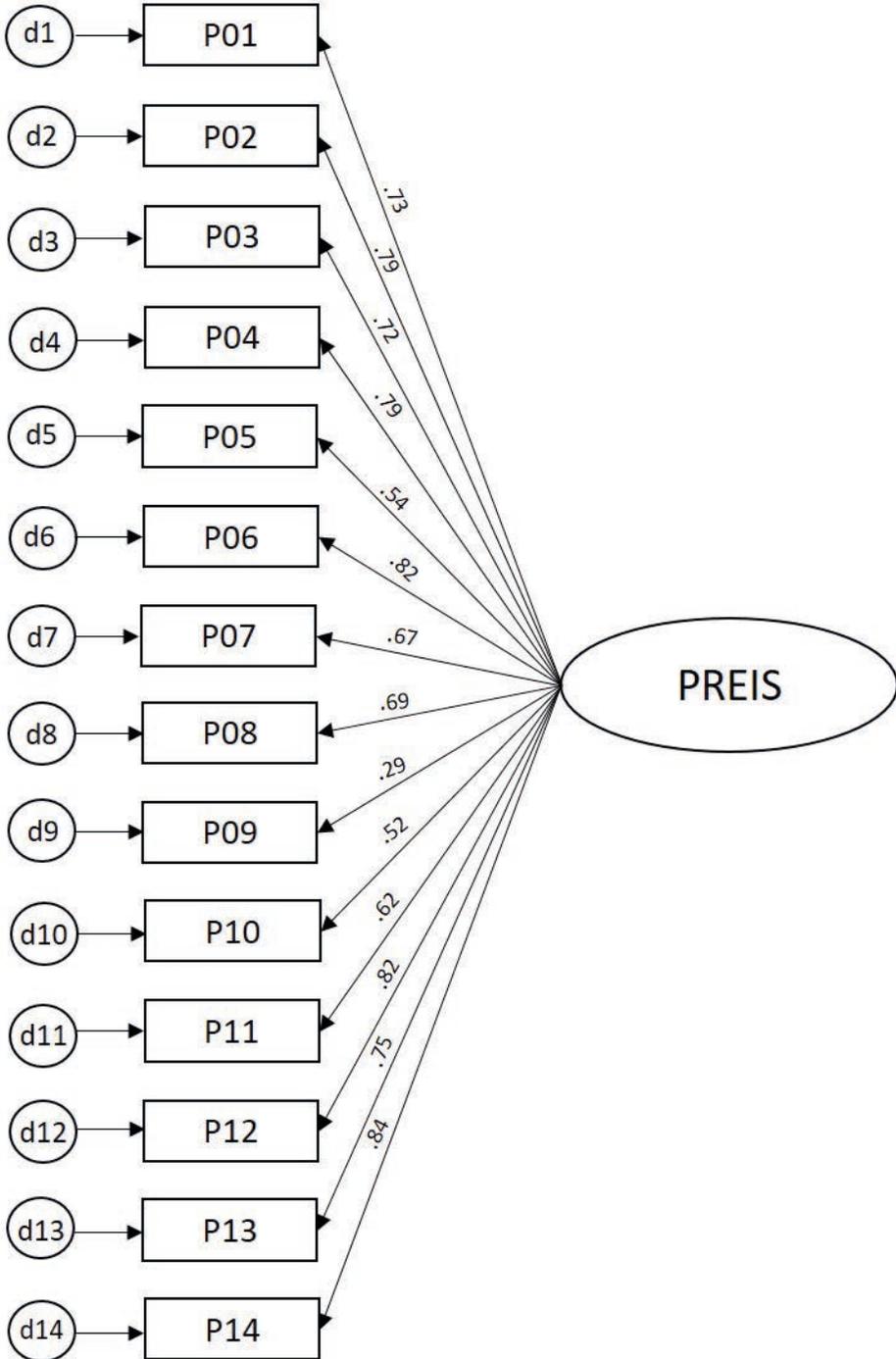


Abbildung 3. Pfadmodell PREIS-E

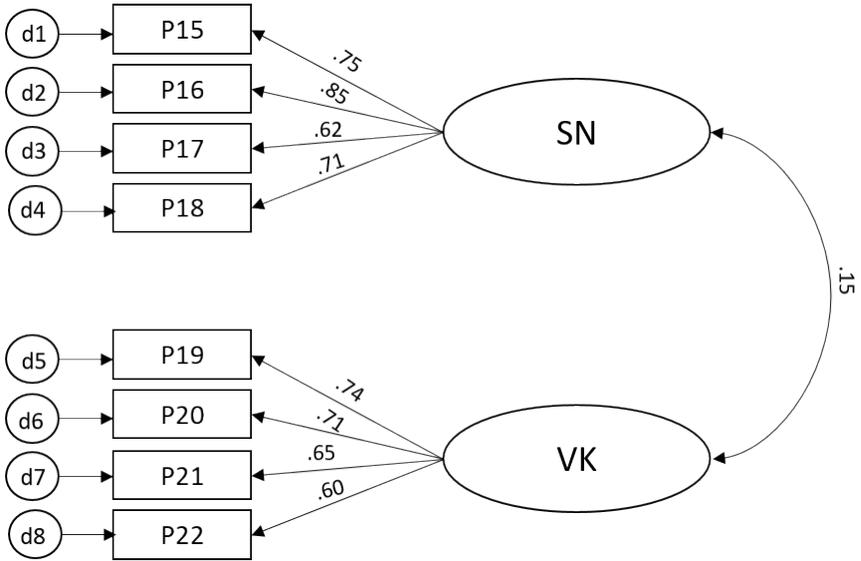


Abbildung 4. Pfadmodell FEDI

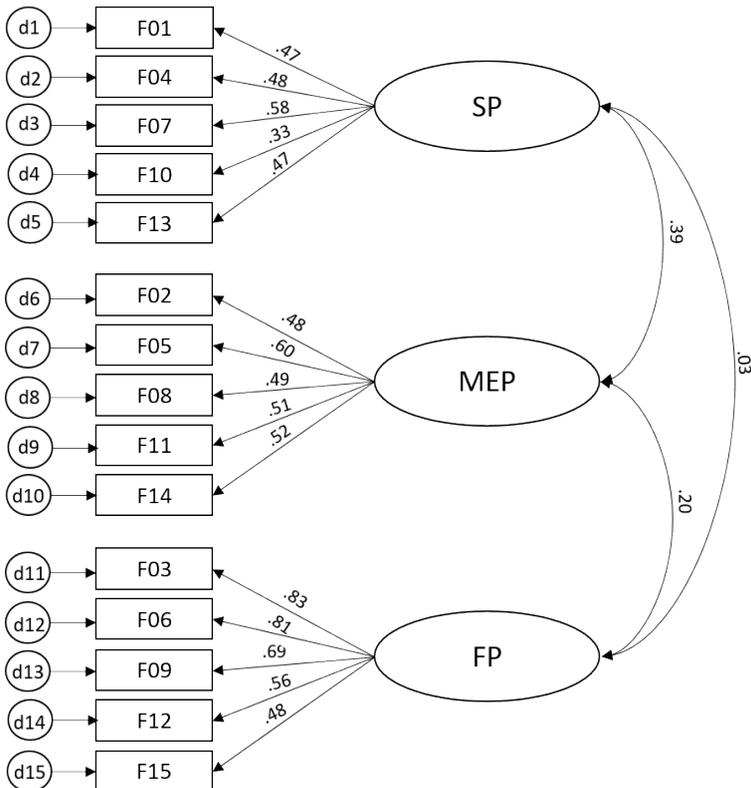


Tabelle 3. Fit-Indizes der konfirmatorischen Faktorenanalysen

Instrument	χ^2	df	χ^2/df	p	CFI	TLI	NFI	RMSEA
PREIS	237.882	78	3.050	<.001	.945	.925	.920	.068
PREIS-E	44.704	19	2.353	<.001	.974	.950	.956	.055
EFI-L	291.979	87	3.356	<.001	.930	.903	.904	.072
FEDI	254.083	87	2.920	<.001	.864	.813	.812	.060

PREIS: 1 Faktor (14 Items)

PREIS-E: 2 Faktoren (SN: 4 Items, VK: 4 Items)

EFI-L: 3 Faktoren (FF: 6 Items, PB: 5 Items, SI: 4 Items)

FEDI: 3 Faktoren (SP: 5 Items, MEP: 5 Items, FP: 5 Items)

Hinweis: Es handelt sich bei allen Modellen um jeweils eigenständige Modelle.

den Einstellungen zu Inklusion zeigen sich keine bis moderate Korrelationen (PREIS: $r = .02-.30$; EFI-L: $r = -.01-.26$), während es keine bzw. schwach negative Zusammenhänge zur wahrgenommenen Verhaltenskontrolle gibt ($r = -.08-.06$). Alle Ergebnisse können Tabelle 2 entnommen werden.

Prüfung der Messinvarianz nach Lehramtsstudiengang

Tabelle 4 zeigt die Ergebnisse der Messinvarianzprüfung. Bei allen Skalen ist der Modellfit unter den Bedingungen der konfiguralen Invarianz ähnlich gut wie bei der initialen CFA. Die Skala PREIS erfüllt die Kriterien für strikte Invarianz, während die PREIS-E lediglich bis zur starken Invarianz ein eindeutiges Ergebnis zeigt. Die Veränderung im RMSEA deutet auch auf strikte Invarianz hin, während der CFI auf Non-Invarianz hindeutet. Bei EFI-L und FEDI deuten die Veränderungen im RMSEA auf strikte Invarianz hin. Die Veränderungen im CFI deuten unter Annahme des strengeren Cut-Offs von 0,01 auf schwache Invarianz bei EFI-L und keine Invarianz bei FEDI hin. Unter Annahme des weniger strengen Kriteriums von 0,02 im CFI ist bei beiden Instrumenten auch die strikte Invarianz erfüllt. Hervorzuheben ist, dass trotz der hohen Teststärke teilweise sogar die χ^2 -Differenz zwischen den Messinvarianzstufen nicht statistisch signifikant ist.

Diskussion

Im Hinblick auf die Reliabilität lässt sich bei allen untersuchten Skalen und Subskalen größtenteils eine akzeptable bis sehr gute interne Konsistenz feststellen, die sich in den Fällen der PREIS und des EFI-L auch in akzeptablen bis guten Retest-Reliabilitäten widerspiegelt. Der FEDI weist jedoch in einer Subskala eine unzureichende interne Konsistenz auf, sowie insgesamt eine schwache Retest-Reliabilität, die auch bei PREIS-E vorliegt.

Bezogen auf die Replikation der Faktorensstrukturen von EFI-L und PREIS knüpfen die Ergebnisse der vorliegenden Untersuchung an bisherige Forschungsergebnisse an, wobei beide Instrumente eine Schwäche in der Chi-Quadrat-Teststatistik aufweisen. Die Struktur der PREIS-E konnte in der vorliegenden Studie repliziert werden. Sowohl auf Grundlage der Modellfit-Indizes als auch auf Faktoren- und Indikatorebene genügt die Skala den gängigen Kriterien. Der FEDI zeigt im Rahmen dieser Studie einige Schwächen: Insgesamt sieben Indikatoren weisen schwache Ladungen auf. Auch im Gesamtmodell deuten einige Fit-Indizes auf eine schlechte Modellpassung hin. Demnach sollte in weiterführenden Studien die Faktorenstruktur nochmals überprüft werden und ggf. ein Alternativmodell aufgestellt werden.

Hinweise auf eine konvergente Validität lassen sich bei PREIS und EFI-L durch eine hohe Interkorrelation feststellen ($r = .74$), was vorherige Forschungsergebnisse be-

Tabelle 4. Prüfung der Messinvarianz nach Lehramtsstudiengang (Primarstufe, Sekundarstufe I/II, Sonderpädagogik)

Invarianz	df	χ^2	$\Delta\chi^2$	p	RMSEA	Δ RMSEA	CFI	Δ CFI
PREIS								
Konfigural	234	425.74			0.074		0.935	
Schwach	258	447.18	21.44	.613	0.070	-0.004	0.936	0.001
Stark	284	487.10	39.92	.040	0.069	-0.001	0.931	-0.005
Strikt	312	537.69	50.59	.006	0.070	0.000	0.923	-0.008
PREIS-E								
Konfigural	57	84.47			0.057		0.973	
Schwach	69	103.08	18.61	.099	0.057	0.001	0.966	-0.007
Stark	81	114.98	11.90	.454	0.053	-0.005	0.966	0.000
Strikt	97	154.17	39.19	.001	0.063	0.010	0.943	-0.023
EFI-L								
Konfigural	261	515.92			0.081		0.914	
Schwach	285	536.02	20.10	.691	0.077	-0.004	0.915	0.001
Stark	309	612.56	76.54	< .001	0.081	0.004	0.897	-0.018
Strikt	339	686.35	73.79	< .001	0.083	0.002	0.882	-0.015
FEDI								
Konfigural	261	465.20			0.072		0.841	
Schwach	285	510.29	45.09	.006	0.073	0.000	0.824	-0.016
Stark	309	558.34	48.05	.002	0.073	0.001	0.806	-0.019
Strikt	339	597.09	38.75	.132	0.071	-0.002	0.799	-0.007

stätigt (Gorges et al., 2020). Der moderate lineare Zusammenhang zwischen Einstellungen und wahrgenommener Verhaltenskontrolle aus der PREIS-E ($r = .28-.31$) ist im Kontext vorliegender Forschungsergebnisse zur Selbstwirksamkeit erwartungsgemäß und liefert weitere Hinweise auf diskriminante Validität (Yada et al., 2022; Gorges et al., 2020; Schulze et al., 2019; Ruberg & Porsch, 2017). Die hohe Korrelation von wahrgenommener sozialer Norm mit den Einstellungen ($r = .54-.62$) ist an die Ergebnisse einer Studie von Wilson et al. (2016) anschlussfähig, die einen Zusammenhang von $r = .53-.54$ berechnet haben, wobei einschränkend hinzuzufügen ist, dass diesbezüglich in anderen Studien geringere Zusammenhänge berichtet werden (z.B. MacFarlane & Woolfson, 2013; Hellmich et al., 2019). Möglicherweise lassen sich diese Unterschiede mit den unterschiedlichen inhaltlichen Ausrichtungen der verwendeten Messinstrumente erklären: Während in der vorliegenden Studie und bei Wilson et al. (2016) die subjektiv wahrgenommene soziale Norm an Menschen ausgerichtet ist, „die mir wichtig sind“, adressieren die Messinstrumente des Konstrukts bei Hellmich et al. (2019) und MacFarlane und Woolfson (2013) Personen der Schulleitung oder -verwaltung. Darüber hinaus könnte sich der enge Zusammenhang mit den Einstellungen damit erklären lassen, dass eine hohe Unterstützung für Inklusion bei Studienteilnehmenden angegeben wird, wenn die Institution, die die Studie durchführt, anscheinend Inklusion unterstützt (Tempel, 2023). Der Name der Institution, in deren Kontext die vorliegende Studie durchgeführt wurde, beinhaltet einen positiv wahrnehmbaren Bezug zu Inklusion, was eine mögliche Erklärung dieser Korrelation sein könnte („Annelie-Wellensiek-Zentrum für Inklusive Bildung“).

Die geringe Korrelation zwischen subjektiv wahrgenommener sozialer Norm und Verhaltenskontrolle ($r = .10$) lässt sich als weiterer Hinweis für diskriminante Validität der Konstrukte interpretieren.

Die Zusammenhänge zwischen Einstellungen und subjektiven Definitionen von Inklusion fallen im Vergleich zur Validierungsstudie des FED1 geringer aus (Faktoren „schulsystemische Perspektive“, $r = .25-.29$; „menschenrechtsbasiert-ethische Perspektive“, $r = .19-.30$) bzw. konnten nicht repliziert werden (Faktor „Förderungsorientierte Perspektive“, $r = .01-.02$). Da diese Korrelationen deutlich geringer ausfallen als die Zusammenhänge zwischen den Instrumenten zu Einstellungen zu Inklusion ($r = .74$), lassen sich diese Ergebnisse als Hinweis auf diskriminante Validität deuten. Die geringen Korrelationen zur wahrgenommenen Verhaltenskontrolle konnten nur beim Faktor „förderungsorientierte Perspektive“ repliziert werden ($r = .08$).

Zugleich liefern die Daten Hinweise darauf, dass sich die Instrumente je nach herangezogenem Cut-Off-Wert als tendenziell messinvariant zwischen den Lehramtsstudiengängen erweisen und für Gruppenvergleiche prinzipiell infrage kommen. Aufgrund der teilweisen Widersprüchlichkeit der Befunde (insbesondere zwischen *RMSEA* und *CFI*) sowie der teilweise kleinen Teilstichproben in einzelnen Lehrämtern ($n = 100-214$) sollten die Befunde zur Messinvarianz nicht überbewertet, sondern eher als Aufgabe für weitere Replikationsstudien aufgefasst werden.

Die vorliegende Studie weist einige Limitationen auf: Erstens lagen die beiden Messzeitpunkte am Anfang und am Ende des Semesters. Da die Studierenden über das Semester Studieninhalte zu den Themen Inklusion und inklusiver Unterricht erhalten haben, könnte die unzureichende Retest-Reliabilität auch auf Lerneffekte zurückzuführen sein, sodass eine wahre Merkmalsveränderung vorliegen könnte. Zweitens liegt ein möglicher Selektionsbias in der Stichprobe vor. Teile der Stichprobe stammen aus freiwillig besuchten Wahlveranstaltungen, andere Teile wiederum aus Pflichtmodulen. Der Datensatz bietet jedoch keine Möglichkeit, zwischen Pflicht-, Wahlpflicht- und Wahlveranstaltungen zu

unterscheiden. Daher kann eine mögliche Verzerrung hier nicht detaillierter analysiert werden.

Zusammenfassend können zum Einsatz der Skalen in zukünftigen Studien folgende Hinweise gegeben werden: Zur Messung von Einstellungen zu Inklusion können die beiden Instrumente EFI-L und PREIS genutzt werden, wobei für die Wahl der Instrumente folgende Entscheidungsgrößen herangezogen werden könnten: Sollte der inhaltliche Fokus auf dem konkreten inklusiven Unterricht liegen, eignet sich EFI-L besonders, wohingegen die PREIS sich besonders auf ein inklusives Schulsystem bezieht. Beide Instrumente können für angehende Lehrkräfte eingesetzt werden, wobei EFI-L explizit für den Einsatz bei Lehrkräften entwickelt wurde und PREIS eine Professionsunabhängigkeit für sich in Anspruch nimmt. Die vorliegende Studie kann jedoch keine Aussage zur Verwendung in anderen Berufsgruppen oder der Allgemeinbevölkerung machen. Die PREIS-E eignet sich gut zur Prüfung der Konstrukte der wahrgenommenen Verhaltenskontrolle und wahrgenommenen sozialen Norm bezogen auf das inklusive Schulsystem. Insbesondere die Reliabilität dieser Skala sollte in fortführenden Studien allerdings weiter geprüft werden. Bezüglich des FEDI sollte zunächst eine Anpassung des Gesamtmodells erfolgen, was beispielsweise durch eine explorative Faktorenanalyse im Rahmen einer weiteren Datenerhebung erfolgen könnte.

Literatur

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- 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. (Hrsg.). (2015). *Fortgeschrittene Multivariate Analysemethoden: Eine anwendungsorientierte Einführung* (3., überarbeitete und aktualisierte Auflage). Springer Gabler. <https://doi.org/10.1007/978-3-662-46087-0>
- Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5
- 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>
- Eagly, A. H. & Chaiken, S. (2010). *The psychology of attitudes* [Nachdr.]. Wadsworth Cengage Learning.
- Egener, L., Scheer, D., Laubenstein, D. & Melzer, C. (2019). *Fragebogen zur Erfassung subjektiver Definitionen von Inklusion (FEDI)*. <https://doi.org/10.13140/RG.2.2.24022.40002>
- Gorges, J., Grumbach, J., Micheel, S. A. & Neumann, P. (2020). Erfassung von Einstellungen zu Inklusion mit SACIE, EFI-L und KIESEL. *Diagnostica*, 1–11. <https://doi.org/10.1026/0012-1924/a000253>
- Grosche, M. & Lüke, T. (2020). Vier Vorschläge zur Verortung quantitativer Forschungsergebnisse über schulische Inklusion im internationalen Inklusionsdiskurs. In C. Gresch, P. Kuhl, M. Grosche, C. Sälzer & P. Stanat (Hrsg.), *Schüler*innen mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen* (S. 29–54). Springer VS. https://doi.org/10.1007/978-3-658-27608-9_2
- Hellmich, F., Löper, M. F. & Görel, G. (2019). The role of primary school teachers' attitudes and self-efficacy beliefs for everyday practices in inclusive classrooms – a study on the verification of the 'Theory of Planned Behaviour'. *Journal of Research in Special Educational Needs*, 19(S1), 36–48. <https://doi.org/10.1111/1471-3802.12476>

- IBM Corp. (2022a). *IBM SPSS AMOS, Version 29* [Computer software].
- IBM Corp. (2022b). *IBM SPSS Statistics for Windows, Version 29.0* [Computer software]. Armonk, NY.
- Jonas, K., Stroebe, W. & Hewstone, M. (Hrsg.). (2014). *Springer-Lehrbuch. Sozialpsychologie* (M. Reiss & K. Jonas, Übers.) (6. Aufl. 2014). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-65297-8>
- Krischler, M., Powell, J. J. W. & Pit-Ten Cate, I. M. (2019). What is meant by inclusion? On the effects of different definitions on attitudes toward inclusive education. *European Journal of Special Needs Education, 34*(5), 632–648. <https://doi.org/10.1080/08856257.2019.1580837>
- Limesurvey GmbH. (2023). *LimeSurvey: An Open Source survey tool* [Computer software]. Hamburg. <http://www.limesurvey.org>
- 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. (2019). *PREIS-E. Erweiterung der PREIS um die wahrgenommene soziale Norm und die wahrgenommene Verhaltenskontrolle*. <https://doi.org/10.6084/m9.figshare.8948780>
- MacFarlane, K. & Woolfson, L. M. (2013). Teacher attitudes and behavior toward the inclusion of children with social, emotional and behavioral difficulties in mainstream schools: An application of the theory of planned behavior. *Teaching and Teacher Education, 29*, 46–52. <https://doi.org/10.1016/j.tate.2012.08.006>
- Makel, M. C. & Plucker, J. A. (2014). Facts Are More Important Than Novelty. *Educational Researcher, 43*(6), 304–316. <https://doi.org/10.3102/0013189X14545513>
- Mandl, H. & Huber, G. L. (1983). Subjektive Theorien von Lehrern. *Psychologie in Erziehung und Unterricht, 30*(2), 98–112.
- Moosbrugger, H. & Kelava, A. (Hrsg.). (2020). *Lehrbuch. Testtheorie und Fragebogenkonstruktion*. Springer. <https://doi.org/10.1007/978-3-642-20072-4>
- Plucker, J. A. & Makel, M. C. (2021). Replication is important for educational psychology: Recent developments and key issues. *Educational Psychologist, 56*(2), 90–100. <https://doi.org/10.1080/00461520.2021.1895796>
- Przibilla, B., Linderkamp, F. & Krämer, P. (2018). Subjektive Definitionen von Lehrkräften zu Inklusion – eine explorative Studie. *Empirische Sonderpädagogik, 10*(3), 232–247. <https://doi.org/10.25656/01:16596>
- R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rosseel, Y. (2023). *Lavaan: Latent Variable Analysis* [R package, version 0.6-16]. <https://cran.r-project.org/package=lavaan>
- Ruberg, C. & Porsch, R. (2017). Einstellungen von Lehramtsstudierenden und Lehrkräften zur schulischen Inklusion. Ein systematisches Review deutschsprachiger Forschungsarbeiten. *Zeitschrift für Pädagogik, 63*(3), 393–415. <https://doi.org/10.25656/01:18583>
- Rutkowski, L., & Svetina, D. (2014). Assessing the Hypothesis of Measurement Invariance in the Context of Large-Scale International Surveys. *Educational and Psychological Measurement, 74*(1), 31–57. <https://doi.org/10.1177/0013164413498257>
- Scheer, D. & Melzer, C. (in Vorb.). Messung subjektiver Definitionen von und Einstellungen zu Inklusion bei pädagogischen Fachkräften in Kindertagesstätten.
- Scheer, D., Egener, L., Laubenstein, D. & Melzer, C. (2023). Development and psychometric evaluation of an instrument measuring subjective definitions of inclusion (FEDI). *International Journal of Inclusive Education, 27*(4), 472–492. <https://doi.org/10.1080/13603116.2020.1853256>
- Schmidt, S. (2009). Shall we Really do it Again? The Powerful Concept of Replication is Neglected in the Social Sciences. *Review of General Psychology, 13*(2), 90–100. <https://doi.org/10.1037/a0015108>
- 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. <http://doi.org/10.13140/RG.2.2.16769.10088>
- Schwab, S., & Helm, C. (2015). Überprüfung von Messinvarianz mittels CFA und DIF-Analysen. *Empirische Sonderpädagogik, 7*(3), 175–193. <https://doi.org/10.25656/01:11380>
- Seifried, S. & Heyl, V. (2016). Konstruktion und Validierung eines Einstellungsfragebogens zu

Inklusion für Lehrkräfte (EFI-L). *Empirische Sonderpädagogik*, 8(1), 22–35. <https://doi.org/10.25656/01:11852>

Tempel, T. (2023). Asking about inclusion: question order and social desirability influence measures of attitudes towards inclusive education. *European Journal of Special Needs Education*, 38(6), 909–915. <https://doi.org/10.1080/08856257.2023.2162666>

Wilson, C., Woolfson, L. M., Durkin, K. & Eliott, 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.2021.103521>

Autorinnen- und Autorenhinweis

-  Christina Mechler
<https://orcid.org/0000-0003-4107-2943>
-  David Scheer
<https://orcid.org/0000-0002-0534-7869>
-  Vera Heyl
<https://orcid.org/0009-0005-9850-8439>

Korrespondenzadresse:

Christina Mechler
Eugen-Neter-Schule
Alter Frankfurter Weg 30
68307 Mannheim
mechlerc@ph-heidelberg.de

Erstmals eingereicht: 30.06.2023
Überarbeitung eingereicht: 14.03.2024
Angenommen: 15.04.2024

Offene Daten	Die Daten sind verfügbar unter: https://doi.org/10.17605/OSF.IO/X73JE
Offener Code	Soweit uns bekannt ist, gibt es keine Möglichkeit, einen Code über SPSS AMOS zu erstellen. Wir können deshalb nichts zur Verfügung stellen. Soweit R für Analysen genutzt wurde, steht die Syntax hier zur Verfügung: https://doi.org/10.17605/OSF.IO/X73JE
Offene Materialien	Die eingesetzten Messinstrumente sind frei zugänglich unter den im Manuskript angegebenen Quellen.
Präregistrierung	Nein
Votum Ethikkommission	Ja, die Studie wurde von der gemeinsamen Ethikkommission der Pädagogischen Hochschule Heidelberg und der SRH Hochschule Heidelberg geprüft. Ein positives Ethikvotum vom 26.7.2021 liegt vor.
Finanzielle und weitere sachliche Unterstützung	keine Angabe
Autorenschaft	CM, DS und VH haben das Manuskript geschrieben. CM, DS und VH haben die Studie geplant. CM hat Daten erhoben, DS und VH haben bei diesen Schritten beraten, CM und DS haben die Daten analysiert.