

*Empirische Sonderpädagogik*, 2022, Nr. 2, S. 190-207  
ISSN 1869-4845 (Print) · ISSN 1869-4934 (Internet)

## Erfassung des Klassenklimas im inklusiven Unterricht – Konstruktvalidität und Eignung für Kinder mit und ohne sonderpädagogischen Förderbedarfen

*Stefan Blumenthal<sup>a</sup>, Michael Schurig<sup>b</sup> & Yvonne Blumenthal<sup>a</sup>*

<sup>a</sup> Universität Rostock

<sup>b</sup> Technische Universität Dortmund

### Zusammenfassung

Studien belegen, dass Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf bedeutsam negativere Selbsteinschätzungen des Klassenklimas im inklusiven Unterricht im Vergleich zu ihren Peers beschreiben. Insofern ist die reliable Erfassung des wahrgenommenen Klassenklimas von hoher pädagogischer Bedeutung.

Die vorliegende Studie untersucht beispielhaft, inwieweit sich drei Skalen des FEES 3-4 (Rauer & Schuck, 2003) zur Erfassung des sozialen Beziehungsgeflechts in inklusiven Schulen eignen. Grundlage der Analysen bilden die Daten von 925 Grundschülerinnen und Grundschulern, die jeweils am Ende des Schuljahres der Klassenstufen 3 und 4 befragt wurden. Anhand konfirmatorischer Faktorenanalysen wurde deutlich, dass die angenommene Struktur der Skalen angepasst werden muss. Insbesondere evozieren negativ formulierte Items problematische Antworttendenzen, die beim Einsatz der Skalen beachtet werden müssen. Unter Berücksichtigung weicher Kriterien konnte die Testfairness zwischen Kindern mit und ohne Förderbedarfen sowie zwischen Jungen und Mädchen nachgewiesen werden. Deutlich werden jedoch Unterschiede in den Datenstrukturen zwischen den Messzeitpunkten, die auf einen Alterseffekt über das Schuljahr hindeuten. Im Rahmen von Mehrebenenstrukturgleichungsmodellen zeigten sich hohe Variabilitäten zwischen den Klassen sowie signifikant niedrigere Klimaerfahrungen bei Kindern mit Förderbedarfen.

*Schlüsselwörter:* Klassenklima, sonderpädagogischer Förderbedarf, inklusiver Unterricht, Grundschule, Längsschnittstudie

## Assessment of classroom climate in inclusive classes - Construct validity and suitability for students with and without special educational needs

### Abstract

Studies show that students with special educational needs describe significantly more negative self-assessments of the classroom climate in inclusive class rooms compared to their peers. In this respect, the reliable measurement of the perceived class climate is of high pedagogical importance.

The present study examines by way of example the extent to which three scales of the FEES 3-4 (Rauer & Schuck, 2003) are suitable for recording the social relationship network in inclusive schools. The analyses are based on data from 925 elementary school students who were surveyed at the end of each school year in grades 3 and 4. Based on confirmatory factor analyses, it became clear that the assumed structure of the scales needs to be adjusted. In particular, negatively worded items evoke problematic response tendencies that must be taken into account when using the scales. Considering soft criteria, test fairness between students with and without special needs as well as between boys and girls could be demonstrated. However, differences in the data structures between the measurement time points became clear, indicating an age effect across the school year. Multilevel structural equation models revealed high between-class variability and significantly lower climate experiences among students with special educational needs.

*Keywords:* Class climate, special education needs, inclusive education, elementary school, longitudinal study

### Einleitung

Das soziale Klima in der Klasse spielt eine tragende Rolle in der schulischen Entwicklung (Beld et al., 2019; Thapa' et al., 2013). Im Kontext inklusiver Schulentwicklungsprozesse besteht ein gesteigertes pädagogisches Interesse an einem positiven Klassenklima, welches zugleich als Merkmal eines qualitativ hochwertigen Unterrichts sowie als Zielkriterium eines gemeinsamen Unterrichts aufgefasst wird (Alansari & Rubie-Davies, 2020; Hascher, 2017). Die im Kontext schulischer Inklusion erhöhte Heterogenität der Schülerschaft in Bezug auf unterschiedliche Merkmale (z. B. unterschiedliche Leistungsniveaus oder emotional-soziale Kompetenzen) ist jedoch als ernstzunehmende Hürde zur Realisierung eines positiven Klassenklimas für alle Schülerinnen und Schüler zu sehen. So deuten bspw. die Ergebnisse von Blumenthal &

Blumenthal (2021) oder Crede et al. (2019) an, dass Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf (SPF) bedeutsam negativere Selbsteinschätzungen des Klassenklimas im inklusiven Unterricht im Vergleich zu ihren Peers beschreiben, analog fällt die Befundlage in Hinblick auf die soziale Eingebundenheit dieser Kinder aus Sicht der Peers aus (z. B. Blumenthal & Blumenthal, 2021; Huber & Wilbert, 2012). Auch die Lehrkraft-Kind-Beziehung, als ein weiterer Aspekt des Klimas in der Klasse, fällt bei vorliegendem SPF häufig geringer aus (z. B. Blumenthal & Blumenthal, 2021).

In Verbindung mit Befunden, die zeigen, dass a) Lehrkräfte erfolgreich pädagogischen Einfluss auf das Beziehungsgeflecht in der Klasse haben, sofern sie um die sozialen Verbindungen der Schülerinnen und Schüler wissen (Mamas et al. 2019; Schurig et al., 2016) sowie b) das durch die Schülerinnen und Schüler bewertete Klassenklima

über die Zeit abnimmt (z. B. Blumenthal & Blumenthal, 2021), stellt sich die Frage nach der Kompetenz der Lehrkräfte, einzelne Aspekte des Klassenklimas korrekt einschätzen zu können. So zeigen beispielsweise die Studien von Köb und Janz (2021) oder Wilbert et al. (2020) auf, dass Lehrkräfte die soziale Stellung von Schülerinnen und Schülern mit SPF nicht sicher einschätzen können.

Diese Befunde unterstreichen die Bedeutung einer systematischen Erfassung des Klassenklimas im inklusiven Unterricht. Als Methode der Wahl werden Befragungen der Schülerinnen und Schüler selbst angesehen, insbesondere mit standardisierten Verfahren (Eder, 2018). Ein aktuelles und häufig in Forschung und Praxis genutztes Verfahren ist der Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern (FEES) von Rauer und Schuck (2003). Am Beispiel des FEES ist es das Ziel der vorliegenden Studie, die Eignung zur Erfassung des Klassenklimas (hier verstanden als Beziehungsgeflecht auf Kind-Kind- sowie Lehrkraft-Kind-Ebene) in inklusiven Klassen zu analysieren. Besondere Relevanz hat dabei die Eignung für die Zielgruppe von Schülerinnen und Schülern mit SPF.

### **Klassenklima als bedeutende Variable im inklusiven Unterricht**

Trotz verstärkter Bemühungen um Fragen nach dem Konstrukt und dessen Auswirkungen auf die schulische Entwicklung von Kindern, fehlt es bislang an einer konsensfähigen Auffassung zum Klimabegriff (Eder, 2018; Grewe, 2003). Die Komplexität des Konstruktes wird in der Begriffsexplikation von Eder (1996, 2018) deutlich. Er unterscheidet drei Typen sowie drei Verwendungsweisen des Klimas in der Klasse. Als Klimatypen müssen demnach das individuelle Klima (subjektiv wahrgenommenes Klima), das aggregierte Klima (durchschnittliche Klimawahrnehmung einer Gruppe)

sowie das kollektive Klima (ähnliche Umweltwahrnehmungen in der Gruppe aufgrund von Kommunikation und Interaktion) unterschieden werden. Fend (1977) postuliert das schulbezogene Klima im Spiegel der Bereiche Inhalt (z. B. Erwartungen und Wertestrukturen in der Schule), Interaktion (Umgang von Schülerinnen und Schülern und Lehrkräften) und Beziehungen (Beziehungen unter den Kindern sowie zwischen den von Schülerinnen und Schülern und Lehrern). Hierbei werden Überschneidungen des Klassenklimas zu weiteren schulisch relevanten Konstrukten deutlich. Als Korrelat des schulischen Wohlbefindens und der Schülergesundheit beschreibt das soziale Klima in der Schule bzw. Klasse nicht nur die Erlebensqualität sozialer Beziehungen zur Lehrkraft und zu den Peers im Kontext schulischer Anforderungen und Erwartungen, sondern ebenso wie sehr sich ein einzelnes Kind angenommen und als Teil der Gruppe fühlt, im Sinne der sozialen Partizipation in der Klasse (Hascher, 2017; Krawinkel et al., 2017; Rauer & Schuck, 2003; Thapa et al., 2013).

Die vorliegende Studie bezieht sich in erster Linie auf das subjektive Beziehungsgeflecht zwischen der Gesamtheit der Akteure in einer Schulklasse, also die Beziehungen und Interaktionen zwischen Lehrkräften sowie Schülerinnen und Schülern als auch der Schülerinnen und Schüler untereinander (Roorda et al., 2011).

### **Empirische Befunde zum Klassenklima bei Kindern mit und ohne SPF**

Auch die empirische Befundlage zum Konstrukt Klassenklima ist schwer interpretierbar (vgl. auch Eder, 2018; Grewe, 2003). In neueren Arbeiten werden zwar vermehrt emotional-soziale Aspekte im Rahmen der schulischen Entwicklung von Schülerinnen und Schülern untersucht, jedoch eher bezogen auf den internationalen Raum und häufig im Querschnitt (z. B. Rathmann et al., 2018). Einstimmig wird ein positiver Zu-

sammenhang zwischen dem erlebten Klima in der Schule bzw. Klasse und der schulischen Motivation sowie der schulischen Leistungen postuliert (Crede & Niehorster, 2012). Ebenso konnten Auswirkungen auf das persönliche Empfinden (z. B. Schulzufriedenheit, schulisches Wohlbefinden, Anstrengungsbereitschaft) und auf die Persönlichkeitsentwicklung (im Sinne einer positiven emotional-sozialen sowie Moralentwicklung, eines positiven Selbstwertgefühls) festgehalten werden (Morin et al., 2014).

Auf der anderen Seite sind negative Klimawahrnehmungen der Schülerinnen und Schüler (im Sinne hoher Anforderungen und Erwartungen der Lehrkraft oder mangelnder Unterstützung) häufig mit psychosomatischen Beschwerden oder Verhaltensauffälligkeiten assoziiert (Rathmann et al., 2018).

Merkmale auf Klassenebene werden als bedeutende Einflussgrößen auf das Klassenklima gesehen (vgl. auch Kausalmodell von Eder, 1996). So wurde eine hohe Variabilität des Klimas zwischen unterschiedlichen Klassen festgestellt (z. B. Blumenthal & Blumenthal, 2021; Rauer & Schuck, 2003; Spörer et al., 2015).

Entsprechend sollen empirische Befunde zum Einfluss der Beziehung zwischen der Lehrkraft und den Schülerinnen und Schülern sowie der Beziehung der Kinder untereinander im Sinne klassenbezogener Indikatoren ausgeführt werden.

### **Die Bedeutung der Lehrkraft-Kind-Beziehung**

Verschiedene Forschungsbefunde verdeutlichen die Rolle einer positiven Lehrkraft-Kind-Beziehung für den schulischen Erfolg sowie für die sozialen Interaktionsprozesse innerhalb einer Klasse (Blatchford et al., 2005; Roorda et al., 2011). Ebenfalls erweist sich die erlebte Gerechtigkeit durch die Lehrkraft als prädiktiv für das schulische Wohlbefinden der Kinder (Aldrich et

al., 2016; Hascher & Hagenauer, 2018). Interaktionen von Kindern mit der Lehrkraft finden jedoch im Rahmen einer asymmetrischen Beziehung statt und sind demnach nicht per se konfliktfrei. Entsprechend korrelieren ungünstige Beziehungen zwischen Lehrkräften sowie Schülerinnen und Schüler mit geringeren schulischen Leistungen bzw. einem geminderten Wohlbefinden (Hascher & Hagenauer, 2018). Diese negativen Folgen fallen besonders bedeutsam für Kinder mit SPF ins Gewicht (Grewe, 2003). Insgesamt scheint sich ein SPF ungünstig auf die Beziehungen zwischen Lehrkräften und Kindern auszuwirken. Dies ist insbesondere dahingehend als problematisch zu bewerten, da in der Lehrkraft-Kind-Beziehung als ein bedeutsamer Prädiktor für die Beziehungen der Klassenmitglieder untereinander zu sehen ist (Blatchford et al., 2005), wobei eine negative Haltung der Lehrkraft dem Kind gegenüber einflussstärker ist als eine positive (Huber et al., 2015).

### **Die Bedeutung der Beziehungen zwischen Schülerinnen und Schülern**

Das Beziehungsgeflecht innerhalb einer Klasse wird als Indikator für den Zusammenhalt bzw. die Zugehörigkeit der Schülerinnen und Schüler innerhalb einer Klasse angesehen (synonym Klassenkohäsion, vgl. Begert & Müller, 2019) und wird häufig mit dem sozialen Klima innerhalb der Klasse in Zusammenhang gebracht (Spörer et al., 2015). Einer hohen Kohäsion der Klasse wird eine protektive Wirkung gegenüber Unterrichtsstörungen sowie Gewalt in der Klasse zugeschrieben (Begert & Müller, 2019).

Die soziale Akzeptanz innerhalb der Klasse wird von individuenbezogenen Aspekten beeinflusst, bspw. schätzen Jungen ihre soziale Einbindung innerhalb der Klasse häufig geringer ein als Mädchen (Krawinkel et al., 2017). Eine weitere wichtige und oft untersuchte Einflussgröße auf die soziale Akzeptanz im schulischen Setting bildet ein

SPF des Kindes. So deutet eine Vielzahl an Forschungsergebnissen an, dass Schülerinnen und Schüler mit SPF in inklusiven Klassen der erhöhten Gefahr sozialer Ablehnung ausgesetzt sind, weniger Freundschaftsbeziehungen führen und sich insgesamt eher einsam fühlen als Kinder ohne SPF (Henke et al., 2017; Hoffmann et al., 2020).

Neben schülerspezifischen Einflussfaktoren spielen weitere, klassenbezogene Aspekte eine Rolle (Crede et al., 2019; Huber & Wilbert, 2012).

## Fragestellung

Die Ausführungen zu Facetten des schulischen Klimas aus den vorangegangenen Abschnitten weisen auf inhaltlicher Perspektive darauf hin, dass mittels des FEES 3-4 (Rauer & Schuck, 2003) Einschätzungen des selbstempfundenen Klassenklimas vorgenommen werden können. Der FEES 3-4 deckt die relevanten Konstrukte des Gefühls des Angenommenseins, des Klassenklimas und der sozialen Integration ab und wurde in mehreren der angeführten Forschungsarbeiten verwendet. Ergänzend wird in diesem Beitrag der Frage nach der Einsetzbarkeit in inklusiven Grundschulklassen aus psychometrischer Sicht nachgegangen. Aus der zuvor dargelegten empirischen Befundlage lassen sich hierzu folgende Fokuspunkte und Arbeitshypothesen für die vorliegende Studie ableiten:

- Die psychometrische Güte der ausgewählten Skalen muss gewährleistet sein. Als Teilkonstrukte eines bereits etablierten Testinstruments sind hier ergänzende Analysen in Hinblick auf die Konstruktvalidität der Skalen notwendig. Es wird in diesem Kontext geprüft, inwieweit die Konstrukte sich auf einer latenten Ebene modellieren lassen, um Modellannahmen explizit testen zu können.
- Die Testfairness zwischen verschiedenen Kindergruppen muss gewährleistet sein. Es ist davon auszugehen, dass Jungen ein geringeres Gefühl der sozialen

Einbindung in der Klasse aufweisen, gleiches gilt für Kinder mit SPF.

- Aufgrund der hohen Variabilität auf Klassenebene, sind mögliche Klasseneffekte zu erwarten und methodologisch zu berücksichtigen. Der FEES adressiert sowohl die individuelle als auch die Klassenebene, so dass angenommen wird, dass substanzielle Varianzanteile auf der Klassenebene berücksichtigt werden müssen, um mögliche ökonomische Fehlschlüsse zu berücksichtigen.

## Methode

### Stichprobe und Vorgehen

Als Datenbasis dienten die Befragungsergebnisse von zwei Grundschulkohorten in Mecklenburg-Vorpommern. Kohorte 1 umfasste alle Schülerinnen und Schüler einer dritten Klasse der öffentlichen Grundschulen der Stadt Stralsund. In Kohorte 2 wurden alle Kinder betrachtet, die zeitlich parallel eine dritte Klasse der öffentlichen Schulen der Region Rügen besuchten. Nach einem Schuljahr wurden die Befragungen ein zweites Mal durchgeführt. Die Stichprobe bestand aus insgesamt  $N = 925$  Schülerinnen und Schülern (50.1 % Mädchen) aus 42 Klassen an 20 Schulen, die zum Ende der vierten Klasse durchschnittlich 10;5 Jahre ( $SD = 0;42$ ) alt waren. Insgesamt wurde darunter eine Gruppe von  $n = 178$  Kindern mit unterschiedlichen Förderbedarfen (34.8 % Mädchen) identifiziert. Der SPF wurde beim schulpsychologischen Dienst abgefragt, dem die Zuständigkeit bei der Feststellungsdiagnostik obliegt. Grundlage für die Diagnosestellung waren die in Mecklenburg-Vorpommern geltenden Standards der Diagnostik. Entsprechend umfasst die hier untersuchte Gruppe von Kindern mit SPF ( $n = 178$ ) Schülerinnen und Schüler mit sonderpädagogischen Förderbedarfen in den Bereichen Lernen ( $n = 8$ ), emotional-soziale Entwicklung ( $n = 43$ ), Sprache ( $n = 3$ ) und Hören ( $n = 1$ ) sowie Teilleistungsstörungen

im Umfang einer Leserechtschreibstörung ( $n = 30$ ), einer isolierten Lesestörung ( $n = 11$ ), einer isolierten Rechtschreibstörung ( $n = 22$ ), einer Rechenstörung ( $n = 4$ ) oder einer kombinierten Störung schulischer Fertigkeiten ( $n = 3$ ). Weiterhin umfasst die Gruppe  $n = 47$  Schülerinnen und Schüler mit einer lernzeitverlängernden Maßnahme im Sinne einer Klassenwiederholung sowie  $n = 6$  weitere Kinder, für die zwar ein Antrag zur Feststellung eines Förderbedarfs gestellt, jedoch abgelehnt wurde. Die letzten beiden Subsamples umfassen zwar keine Schülerinnen und Schüler mit offiziellen Förderbedarfen, umschreiben jedoch Kindergruppen mit deutlichen Risiken für die schulische Entwicklung.

Aufgrund von Fehlzeiten oder nicht-auswertbaren Daten ergaben sich einzelne fehlende Werte. Die Missings in den Zielvariablen weisen keine erkennbare Struktur auf und betreffen maximal 8 % der Daten. Mittels Imputation ( $k = 1$ ) wurden einzelne fehlende Werte ersetzt.

## Instrumente

Als Indikatoren des Klassenklimas wurden die Subskalen Gefühl des Angenommenseins (GA), Klassenklima (KK) und Soziale Integration (SI) des FEESS 3-4 (Rauer & Schuck, 2003) eingesetzt.

### Der Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern (FEESS)

Der FEESS 3-4 dient zur Erfassung emotionaler und sozialer Erfahrungen von Grundschulkindern. Bei der Durchführung sollen die Kinder 90 Aussagen auf einer vierstufigen Skala (0 = „stimmt gar nicht“ bis 3 = „stimmt genau“) bewerten. Die Items verteilen sich auf insgesamt sieben Dimensionen (Selbstkonzept der Fähigkeit, soziale Integration, Klassenklima, Schuleinstellung, Lernfreude, Anstrengungsbereitschaft, Gefühl des Angenommenseins). Es werden getrennte Normwerte für das dritte und das vierte Schuljahr aufgeführt. Eine geschlechtsspezifische Norm für den FEESS 3-4 liegt hingegen nicht vor.

**Tabelle 1**  
*Deskriptive Statistik*

Skala	Gruppe	N	M	SD
Alter	Kein SPF	687	10.5	0.40
	SPF	157	10.6	0.50
Geschlecht	Kein SPF	739	0.54	0.50
	SPF	178	0.35	0.48
Gefühl des Angenommenseins Kl. 3	Kein SPF	699	22.7	2.99
	SPF	156	22.7	4.52
Gefühl des Angenommenseins Kl. 4	Kein SPF	708	28.8	7.11
	SPF	165	26.3	7.92
Soziale Integration Kl. 3	Kein SPF	697	25.1	6.60
	SPF	155	22.3	7.78
Soziale Integration Kl. 4	Kein SPF	706	23.7	6.36
	SPF	167	21.1	7.51
Klassenklima Kl. 3	Kein SPF	697	22.5	6.21
	SPF	155	20.4	7.33
Klassenklima Kl. 4	Kein SPF	706	20.2	6.24
	SPF	167	18.6	6.91

## Gefühl des Angenommenseins (GA)

Die Skala GA bezieht sich auf die eigene Wahrnehmung des persönlichen Verhältnisses zur Lehrkraft (Beispielitem „Meine Lehrer mögen mich.“). Die Skala umfasst 13 Items, die im Rahmen der vorliegenden Untersuchung interne Konsistenzen in Höhe von  $\alpha = .86$  sowohl zum Ende der Klasse 3 sowie 4 aufwiesen.

## Klassenklima (KK)

Mit der Skala KK wird mittels elf Items das soziale Miteinander über die wahrgenommenen sozialen Beziehungen in der Klasse erfasst (Beispielitem „Wir lassen nicht alle Kinder mitmachen.“). In der vorliegenden Stichprobe zeigte sich sowohl zum Ende der Klasse 3 als auch zum Ende der Klasse 4 eine als gut einzuschätzende Reliabilität von jeweils  $\alpha = .80$ .

## Soziale Integration (SI)

Die elf Items der Skala SI beziehen sich auf das wahrgenommene Ausmaß des erlebten sozialen Eingebundenseins in der Klasse (Beispielitem „Meine Mitschüler sind nett zu mir.“). In der vorliegenden Stichprobe ergeben sich Werte zur internen Konsistenz in Höhe von  $\alpha = .86$  (Klasse 3) bzw.  $\alpha = .83$  (Klasse 4).

## Unabhängige Variablen und Kovariaten

Als Hintergrundmerkmale wurden sowohl die Klassenstufe und das Geschlecht als auch das Vorliegen eines SPF kontrolliert. Indikatoren für die Klassenzugehörigkeit wurden als weitere Variablen berücksichtigt, um mögliche Varianzen zwischen den Klassen zu kontrollieren. Der Intraklassenkoeffizient (ICC) der manifesten Items liegt zwischen .03 (SI Item 10) und .16 (KK Item 1) und unterstreicht damit die Notwendigkeit der Berücksichtigung beider Ebenen (Hox, 2010).

## Vorgehen bei der Datenauswertung

Zunächst wurden Item- und Skalenstatistiken (Cronbach's  $\alpha$  und McDonald's  $\omega$  sowie Verteilungsstatistiken und Trennschärfen) für die drei beschriebenen Skalen bestimmt, welche jedoch aus Platzgründen nicht in das Manuskript integriert, sondern in das elektronische Supplement ausgelagert wurden.

In der Datenanalyse wurde auf kongenerische Modelle zurückgegriffen, um Annahmen zu Faktorladungen und der Kovarianz multipler latenter Konstrukte explizit prüfen zu können. Zunächst wurde eine mehrdimensionale konfirmatorische Faktorenanalyse (CFA) im Rahmen der Strukturgleichungsmodellierung (SEM; Brown 2015) vorgenommen. Grundsätzlich wurden robuste Maximum-Likelihood-Parameterschätzung angewendet. Um die Dimensionalität zu prüfen, wurden Likelihood-Ratio-Tests (LRT) gegen eindimensionale Modelle durchgeführt.

Weiterhin wurden Analysen zur Messinvarianz (MI) zwischen den Testzeitpunkten, den Geschlechtern und Kindern mit sowie ohne SPF angewandt. Prüfungen der Messinvarianz sind in besonderem Maße dafür geeignet, die Testfairness zwischen Gruppen zu prüfen, da partielle und globale Gleichheitsannahmen quantifiziert werden können. Dabei wurden steigend restriktive Modelle gegeneinander geprüft, um festzustellen, ob das konfigurale Setup, die Faktorladungen und die Itemmittelwerte ausreichend ähnlich sind, um nachfolgende Vergleiche durchzuführen (Cheung & Rensvold, 1998). Eine Vergleichbarkeit der latenten Mittelwerte wurde angenommen, wenn mindestens (schwache) faktorielle Invarianz vorliegt. Geschachtelte Modelle wurden dafür verglichen, um Invarianzgrade zu schätzen. Rutkowski und Svetina (2014) schlagen zusätzlich zur Verwendung von LR-Tests Cut-off-Werte von  $\Delta CFI \leq 0.02$  und  $\Delta RMSEA \leq 0.03$  vor, um das Vorhandensein einer ausreichenden Messinvarianz festzustellen.

Die geschachtelte Struktur der Daten wird berücksichtigt, indem die Klassenebene als zweites Level modelliert wird (z. B. Hox, 2010). Varianzanteile latenter Konstrukte können damit auf individuelle und Gruppenebene parallel verarbeitet werden.

Die Analysen wurden in R (R Core Team 2020) mit den Paketen *lavaan* (Rosseel, 2012) und *semTools* (Jorgensen et al., 2020) durchgeführt. Die Visualisierung der Modelle erfolgte mit *Ωnyx* (von Oertzen et al., 2015).

## Ergebnisse

### Konfirmatorische Faktorenanalysen

Im ersten Schritt wurden die multidimensionalen Modelle jeweils innerhalb der Jahrgangsstufen 3 und 4 geprüft. Das Modell in der Jahrgangsstufe 3 (Modell 1.3) erreichte dabei ausreichende absolute Fit-Statistiken ( $\chi^2 = 1661.74$ ,  $df = 557$ ) und approximative Fit-Statistiken (RMSEA = 0.049; SRMR = 0.055) (zur Interpretation der Anpassungswerte vergleiche Kline, 2015). Die Anpassungsstatistiken gegen ein Basismodell zeigten jedoch inakzeptable Werte (CFI = 0.835; TLI = 0.824). Im Modell liegen mehrere Faktorladungen  $< .5$  vor.

Im Vergleich mit einem g-Faktor Modell (Modell 2.3) zeigte sich das mehrdimensionale Modell mittels eines LRT als angemessener ( $\Delta\chi^2 = 463.74$ ,  $\Delta df = 3$ ,  $p < 0.001$ ).

Das Modell in der Jahrgangsstufe 4 (Modell 1.4) war auch in Bezug auf absolute Fit-Maße nicht akzeptabel ( $\chi^2 = 5004.56$ ,  $df = 557$ ; RMSEA = 0.096; SRMR = 0.138; CFI = 0.574; TLI = 0.545). Der LR-Test gegen ein g-Faktor-Modell (Modell 2.4) bestätigte die grundlegend höhere Angemessenheit des mehrdimensionalen Modells ( $\Delta\chi^2 = 449.30$ ,  $\Delta df = 3$ ,  $p < 0.001$ ).

Die Messinvarianzanalysen zwischen den Klassenstufen 3 und 4 mittels LR-Tests und die Vergleiche der CFI-Werte bestätigten eine Varianz zwischen der Konfiguration der Modelle ( $\Delta\chi^2 = 85.66$ ,  $\Delta df = 32$ ,

$p < 0.001$ ;  $\Delta CFI = 0.261$ ). Somit mussten die Modifikationen der Modelle auf individueller Ebene und getrennt vorgenommen werden.

Dafür wurden die Faktorladungen in den Klassenstufen 3 und 4 getrennt und gemeinsam inspiziert und mit den Formulierungen problematischer Items verglichen. Dabei fiel auf, dass diejenigen Items, die negativ formuliert sind, systematisch geringere Faktorladungen aufwiesen als die Items, die positiv formuliert sind. Daher wurde ein weiterer Faktor definiert, der sich aus allen negativ formulierten Items zusammensetzt. Der zusätzliche Faktor umfasst vier Items aus der Skala SI, sechs Items für die Skala KK und fünf Items für die Skala GA.

Für die Klassenstufe 3 (Modell 3.3:  $\chi^2 = 1406.15$ ,  $df = 539$ ; RMSEA = 0.043; SRMR = 0.050; CFI = 0.872; TLI = 0.861) erreichte der zusätzliche Faktor für die negativ formulierten Items der Dimension KK Signifikanz und Faktorladungen  $> .27$  (vgl. Kline, 2015). Die Faktorladungen im Bereich der Skalen SI und GA erreichten hingegen keine Signifikanz und werden im Folgenden nicht mehr berücksichtigt. In der Klassenstufe 4 (Modell 3.4:  $\chi^2 = 1691.06$ ,  $df = 539$ ; RMSEA = 0.050; SRMR = 0.049; CFI = 0.890; TLI = 0.878) erreichten alle negativ gepolten Items signifikante Werte und standardisierte Faktorladungen  $> .29$ .

Auf der Basis des Verhältnisses von  $\chi^2$  und  $df$  der Modelle sowie der Maße, die auf den Fehlern der Modelle basieren (RMSEA und SRMR) erscheinen beide Modelle annehmbar. Für die folgenden Analysen wurde jeweils für Klassenstufe 3 und 4 das dreidimensionale Modell 3 inklusive des zusätzlichen Faktors verwendet.

### Messinvarianzanalysen

In einem weiteren Schritt wurden die vorliegenden Modelle auf Messinvarianz getrennt für die Klassenstufen in Abhängigkeit des Geschlechts sowie des Vorliegens eines SPF geprüft (vgl. Tabelle 2 und Tabelle 3).

Generell reagieren LR-Tests restriktiver als Cut-off-Werte auf Basis der Anpassungsmaße. Auf Basis der liberaleren Kriterien der Differenzen des CFI und des RMSEA konnte hingegen sowohl in der Klassenstufe 3 schwache als auch in der Klassenstufe 4 starke faktorielle Invarianz angenommen werden. Da in der Klassenstufe 3 keine starke faktorielle Invarianz zwischen den Geschlechtern festgestellt werden konnte, musste angenommen werden, dass nur eine eingeschränkte direkte Vergleichbarkeit innerhalb der Konstrukte zwischen den Geschlechtern vorliegt.

In Bezug auf den SPF konnte bei einer liberalen Betrachtung der Cut-off-Werte in den Klassenstufen 3 und 4 faktorielle Invarianz beobachtet werden. Die Werte, die auf Basis dieser Konstrukte innerhalb der Jahrgänge abgeleitet werden, sind also vergleichbar.

## Mehrebenenmodellierung

Im letzten Schritt wurde zuerst die Klassenebene gemeinsam parallel zur Individualenebene modelliert (Hox, 2010). Dafür wurden sowohl in der dritten als auch in der vierten Klassenstufe jeweils alle 42 Klassen verarbeitet. Die Zahl der Kinder pro Klasse lag dabei zwischen 11 und 29. Schrittweise wurden zudem weitere Variablen ergänzt: Bereits im ersten Schritt wird der zusätzliche Faktor der negativ formulierten Items eingeführt. In einem zweiten Schritt wurden das Geschlecht und der SPF eingeführt (vgl. Abbildung 1).

Das Modell der Klassenstufe 3 zeigte dabei eine diskutable Passung auf Basis des  $\chi^2$  sowie des RMSEA und des SRMR und eine schlechte Passung auf Basis des CFI und des TLI ( $\chi^2 = 2252.56$ ,  $df = 1105$ ; RMSEA = 0.035; SRMR = 0.049; CFI = 0.869; TLI = 0.859), während das Modell der Klassenstufe 4 ( $\chi^2 = 3501.24$ ;  $df = 1096$ ; RMSEA = 0.050; SRMR = 0.052; CFI = 0.847; TLI = 0.834) eine noch weniger annehmbare Passung auf Basis des  $\chi^2$  erreichte.

**Tabelle 2**

*Invarianz in der Klassenstufe 3*

	Invarianz-Niveau	BIC	$\chi^2$	$\Delta\chi^2$	df	$\Delta df$	$p$	CFI	$\Delta CFI$	RMSEA	$\Delta RMSEA$
SPF	konfigural	71723	2548.8	-	1096	-	-	0.863	-	0.046	-
	faktoriell (schwach)	71550	2625.2	67.70	1133	37	0.002	0.862	0.000	0.046	0.000
	faktoriell (stark)	71435	2718.7	93.63	1164	31	<0.001	0.851	0.011	0.047	0.001
Geschlecht	konfigural	71566	2595.9	-	1096	-	-	0.863	-	0.046	-
	faktoriell (schwach)	71392	2670.9	58.48	1133	37	0.014	0.860	0.003	0.046	0.000
	faktoriell (stark)	71330	2817.4	150.96	1164	31	<0.001	0.913	0.070	0.047	0.001

**Tabelle 3***Invarianz in der Klassenstufe 4*

	Invarianz-Niveau	BIC	$\chi^2$	$\Delta\chi^2$	df	$\Delta$ df	$p$	CFI	$\Delta$ CFI	RMSEA	$\Delta$ RMSEA
SPF	konfigural	72059	2888.6	-	1078	-	-	0.882	-	0.053	-
	faktoriell (schwach)	71840	2980.5	78.99	1124	46	0.002	0.879	0.003	0.053	0.000
	faktoriell (stark)	71680	3030.7	50.10	1155	31	0.016	0.877	0.002	0.052	0.001
Geschlecht	konfigural	72038	2977.3	-	1078	-	-	0.877	-	0.054	-
	faktoriell (schwach)	71791	3041.5	50.87	1124	46	0.288	0.877	0.000	0.053	0.001
	faktoriell (stark)	71686	3146.0	104.55	1155	31	<0.001	0.871	0.006	0.053	0.000

Die Varianzanteile der Klassenebene gegenüber der Gesamtvarianz wurden be-

stimmt als

$$ICC = \frac{\text{Varianz}_{\text{between}}}{\text{Varianz}_{\text{within}} + \text{Varianz}_{\text{between}}}$$

und beliefen sich in der Klassenstufe 3 auf  $ICC_{KK} = 20.4\%$ ,  $ICC_{GA} = 5.3\%$  und  $ICC_{SI} = 3.1\%$  und in der Klassenstufe 4 auf  $ICC_{KK} = 28.3\%$ ,  $ICC_{GA} = 1.6\%$  und  $ICC_{SI} = 27.8\%$ . Die Varianzanteile variierten zwischen den Klassenstufen und bezogen sich dabei mit erwartbar hohem Anteil auf die Klassenebene.

In einer finalen Modellanpassung wurden die Kovariaten Geschlecht und SPF auf der Individualebene ergänzt (Abbildung ). Die vollständigen Modelle der Klassenstufe 3 ( $\chi^2 = 2636.95$ ;  $df = 1167$ ;  $RMSEA = 0.039$ ;  $SRMR = 0.050$ ;  $CFI = 0.847$ ;  $TLI = 0.835$ ) und der Klassenstufe 4 ( $\chi^2 = 2459.31$ ;  $df = 1152$ ;  $RMSEA = 0.036$ ;  $SRMR = 0.050$ ;  $CFI = 0.879$ ;  $TLI = 0.869$ ) erreichten erneut eine nur diskutable Anpassungsgüte. Dies wird vor allem in den CFI und TLI deutlich. Die Zusammenhänge zwischen den Skalen und die Effekte der Kovariaten sind der Tabelle 4 bzw. der Tabelle 5 zu entneh-

men. Besondere Aufmerksamkeit muss der Korrelation  $r_{SI-KK}$  gewidmet werden, die die übliche Spannweite standardisierter Effekte (-1:1) übersteigt. Dies kann in multivariaten Analysen auf das Vorliegen starker Korrelationen zwischen Prädiktoren zurückgeführt werden (Multikollinearität; z. B. Campbell & Kenny, 1999).

Die Varianzanteile der Klassenebene liegen in der Klassenstufe 3 bei  $ICC_{KK} = 20.5\%$ ,  $ICC_{GA} = 0.1\%$  und  $ICC_{SI} = 2.6\%$  und in der Klassenstufe 4 bei  $ICC_{KK} = 28.5\%$ ,  $ICC_{GA} = 1.7\%$  und  $ICC_{SI} = 28.1\%$ .

**Tabelle 4**  
Effekte Klassenstufe 3

within		KK	GA	SI	NI	between		
						KK	GA	SI
KK	r	-				KK	r	-
GA	r	0.441***	-			GA	r	0.577
SI	r	0.893***	0.51***	-		SI	r	1.073**
NI	r	-0.792	-0.258*	-0.694**	-			0.553
Geschlecht	β	0.014	0.132*	-0.030	0.060			
SPF	β	-0.101	-0.164***	-0.171***	-0.014			

Anmerkungen. KK = Klassenklima; GA = Gefühl des Angenommenseins; SI = Soziale Integration; NI = Negative Items; \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

**Tabelle 5**  
Effekte Klassenstufe 4

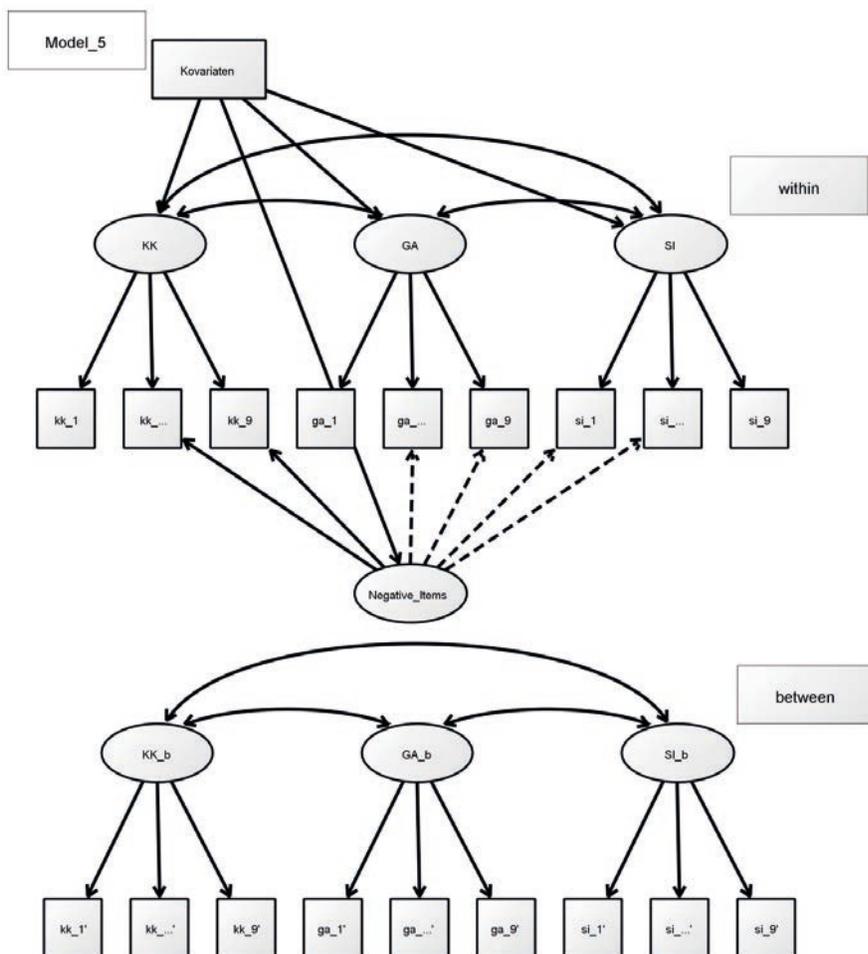
within		KK	GA	SI	NI	between		
						KK	GA	SI
KK	r	-				KK	r	-
GA	r	0.625**	-			GA	r	-0.943
SI	r	0.937*	0.61**	-		SI	r	0.982
NI	r	-0.324	-0.244	-0.324	-			-0.961
Geschlecht	β	0.037	0.093	0.014	0.089*			-
SPF	β	-0.041	-0.094*	-0.128*	-0.094			

Anmerkungen. KK = Klassenklima; GA = Gefühl des Angenommenseins; SI = Soziale Integration; NI = Negative Items; \*  $p < .05$ , \*\*  $p < .01$ , \*\*\*  $p < .001$ .

## Diskussion

Der Ausgangspunkt des vorliegenden Beitrags ist in der hohen Bedeutung des Klassenklimas für die schulische Entwicklung von Schülerinnen und Schülern zu sehen (Beld et al., 2019). Die Erfassung des Klassenklimas sollte, aufgrund eingeschränkter Wahrnehmungen durch die Lehrperson (eher positiv verzerrter Blick), mittels stan-

dardisierter Instrumente erfolgen (Eder, 2018). Für den Grundschulbereich hat sich der FEES (Rauer & Schuck, 2003) als anerkanntes Verfahren zur Einschätzung der emotional-sozialen Schulerfahrungen von Schulkindern erwiesen. Inwieweit der FEES sich jedoch zur Messung der Ausprägung klimabezogener Aspekte in inklusiven Klassen eignet, wurde bislang noch nicht hinreichend untersucht. Dieser Forschungs-



**Abbildung 1**  
 Mehrebenenmodell mit Kovariaten auf Individualebene

frage widmet sich der vorliegende Beitrag. Die gemeinsame Skalenstruktur ließ sich nicht ohne Weiteres abbilden. In einem ersten Schritt wurden die drei Skalen KK, SI und GA als Facetten eines g-Faktors modelliert. Zwar gehen die Autoren des FEES originär nicht von einer gemeinsamen latenten Dimension dieser Skalen aus (Rauer & Schuck, 2003), sondern subsumieren die Skalen KK und SI in der Dimension Sozial-

klima und verstehen das GA als Facette des Schul- und Lernklimas, jedoch gibt es inhaltliche Argumente, die für eine gemeinsame Betrachtung der Skalen sprechen. So wird zum einen die Lehrkraft-Kind-Beziehung als wesentlicher Einflussfaktor für das soziale Klima in der Klasse angesehen (Aldrich et al., 2016; Hascher & Hagenauer, 2018), zum anderen sprechen die in der Skala GA formulierten Items weniger lernbezogene,

sondern eher beziehungsbezogene Aspekte an. Ebenso wurden die postulierten Dimensionen des FEES explorativ entwickelt, jedoch nicht konfirmatorisch geprüft.

Die Annahme eines g-Faktors musste jedoch, aufgrund ungenügender Fit-Statistiken, verworfen werden. Auch die Modellierung der drei untersuchten Skalen als eigenständige Dimensionen erwies sich unter psychometrischen Gesichtspunkten nicht als zielführend. Eine ergänzende explorative Analyse der Faktorenstruktur verwies auf die Notwendigkeit, die Items der drei Skalen jeweils noch unterteilt zu betrachten. Im Ergebnis erwies sich das Modell, das drei separate Dimensionen (KK, SI und GA) und zudem eine weitere Dimension annimmt, die alle negativ formulierten Items aller drei Skalen umfasst, als annehmbar. Die Bedeutung der zusätzlichen Dimension ist anhand mehrerer Annahmen zu erklären. Insgesamt ist eine Verzerrung im Antwortverhalten durch eine ungünstige Itemformulierung (hier doppelte Verneinung) denkbar (Moosbrugger & Kelava 2012). Differenzen der Verteilungen von negativ formulierten Items im Sinne einer tendenziell geringeren mittleren Ausprägung bei etwas höheren Standardabweichungen zeigen sich bereits im Manual des FEES 3-4 (Rauer & Schuck, 2003). Für die Skalen KK und GA ist zu vermuten, dass die negativ formulierten Items (z. B. „Wir lassen nicht alle Kinder mitmachen.“ bzw. „Meine Lehrer sind ein wenig unfreundlich zu mir.“) sozial erwünschte Antworten bei den befragten Kinder im Sinne einer Fremdtäuschung der Lehrkräfte evozieren und in der Skala SI auf Selbsttäuschungsversuche der befragten Kinder hindeuten könnten (Moosbrugger & Kelava, 2012). Die vergleichsweise hohen Itemmittelwerte (bis auf wenige Ausnahmen > 2 bei einer Skalenbreite von 0 bis 3) sprechen darüber hinaus für akquieszente Antworttendenzen (vgl. Bogner & Landrock, 2015).

Die Analysen liefern zudem Hinweise auf eine eingeschränkte Messinvarianz der Skalen. Dies gilt zum einem in Bezug auf

das Geschlecht der befragten Kinder, zum anderen hinsichtlich des Vorhandenseins eines SPF. Zieht man zur Beurteilung liberalere Kriterien heran, ergaben sich zwischen den Geschlechtern sowie zwischen Kinder mit als auch ohne SPF bedingt vergleichbare Datenstrukturen zum Ende der Klasse 3 als auch ein Jahr später.

Die Analysen weisen auf einen signifikanten Einfluss des SPF auf die Ergebnisse in den Skalen GA und SI hin. Diese Befunde korrespondieren mit weiteren Studienergebnissen, wonach die Lehrkraft-Kind-Beziehung (z. B. Blumenthal & Blumenthal, 2021) als auch die soziale Stellung (z. B. Crede et al., 2019; Henke et al., 2017; Hoffmann et al., 2020) negativ von einem vorliegenden SPF beeinflusst wird. Hinsichtlich der Skala KK und dem zusätzlich gebildeten Faktor mit den negativ formulierten Items konnte kein Einfluss des SPF festgestellt werden.

Weiterhin zeichnet sich ein bedeutsamer Alterseffekt vom Ende der Klasse 3 zum Ende der Klasse 4 ab, dies gilt insbesondere für den zusätzlich gebildeten Faktor. Dies ist insofern erwartungskonform, als dass sozial erwünschte Antworten bei jüngeren Kindern seltener sind, die im Gegenzug jedoch häufiger zum Ja-Sagen tendieren (Vogl, 2015). Für die Skala GA zeigten bereits Rauer und Schuck (2003) Unterschiede in den Verteilungen zwischen den Klassenstufen. Auch andere Studienbefunde konnten einen bedeutsamen Rückgang des erlebten Klassenklimas in Abhängigkeit vom Alter festhalten. Dies gilt für die Beziehung zwischen Lehrkräften und den Kindern (Grewe, 2003) sowie für das empfundene Erleben in der Klasse (Gentry et al., 2002). Hieraus lassen sich zwei Annahmen ableiten: a) die eingeschränkten diagnostischen Kompetenzen der Lehrkräfte (eher rosy view, vgl. z. B. Köb & Janz, 2021) führen zu einem mangelnden Problembewusstsein hinsichtlich der sozialen Strukturen in der Klasse und b) es scheint, dass es Lehrkräften schwerfällt, das soziale Beziehungsgefüge in der Klasse aktiv und positiv zu gestalten (Grewe, 2003). Beide Annahmen jeweils

für sich oder in Kombination, stehen in Zusammenhang mit der Variabilität der Befunde zum Klima zwischen den Klassen (z. B. Blumenthal & Blumenthal, 2021; Spörer et al., 2015) die auch im Rahmen der hier vorgelegten Mehrebenenmodellierungen festgehalten werden konnten. Hier fällt vor allem die Differenz in der sozialen Integration zwischen den Klassenstufen auf. Die Werte deuten darauf hin, dass erst in der Klassenstufe 4 ein homogenes Urteil zur sozialen Integration in der Klasse entsteht. Zuvor handelt es sich also um ein Konstrukt, welches vor allem auf der Individualebene verortet werden müsste.

Ein weiteres Problem zeigte sich in der vorliegende Multikollinearität (vgl. Tab. 4 und 5). Das Vorliegen von starken Kovarianzen in Prädiktoren kann in erhöhten Standardfehlern resultieren, welche die Bestimmung der Signifikanz beeinflussen können. Auch können Probleme bei der Schätzung standardisierter Parameter auftreten (Deegan, 1978). Gleichzeitig sind Strukturgleichungen auf der Berücksichtigung und Kontrolle von multikollinearen Annahmen aufgebaut (Kenny, 1979). Deegan (1978) sowie Campbell und Kenny (1999) empfehlen hier keinen mechanischen Ausschluss, sondern eine theoretische Berücksichtigung und gegebenenfalls die Identifikation einer empirischen Unterrepräsentation der Komplexität des Konstrukts. Dies kann zukünftig durch die Verwendung von bedeutsamen, aber für die Analyse wenig relevanten Instrumentalvariablen auf der Klassenebene adressiert werden (Kenny, 1979). Für die vorliegende Analysen ist anzunehmen, dass die Power des Modells auf zweiter Ebene nicht ausreichend gewesen ist, um die Effekte trennscharf auf Signifikanz zu prüfen. Nicht-signifikante Effekte sollten also, trotz ihrer Größe, nicht inhaltlich interpretiert werden, sondern wurden der Transparenz wegen angegeben.

Die geringen Modellanpassungen müssen vor allem auf die inkrementellen Fit-Werte des CFI und des TLI zurückgeführt werden. Da diese besonders sensitiv für ge-

ringe Faktorladungen sind, lässt sich auf Basis der Vergleichsanalysen schlussfolgern, dass die Struktur grundlegend erhalten werden kann, aber die Trennschärfen einzelner Items geringer sind als wünschenswert (Brown, 2015).

Vor dem Hintergrund der Zielsetzung des vorliegenden Beitrags bleibt zu konstatieren, dass Kinder mit SPF dahingehend als gefährdet angesehen werden müssen, als dass ihre sozialen Beziehungen zur Lehrkraft sowie zu den Klassenkameradinnen und -kameraden durch den SPF negativ beeinflusst werden. Eine Befragung hinsichtlich des Klimas in der Klasse kann entsprechend wertvolle Informationen für Lehrkräfte bergen. Die hier dargestellten Befunde zeigen, dass die drei verwendeten Skalen des FEES 3-4 (Rauer & Schuck, 2003) zur Beschreibung des sozialen Beziehungsgeflechts in der Klasse hohe Zusammenhänge aufweisen, sich jedoch mit Einschränkungen als diskrete Konstrukte etablieren lassen. Unter Annahme liberaler Kriterien ist die Testfairness hinsichtlich des Geschlechts und des SPF gegeben. Die negativ formulierten Items der Skalen führen jedoch zu veränderten Antwortstrukturen, was als problematisch zu werten ist, zumal hierbei geschlechtsspezifische Unterschiede in Klasse 4 festgehalten wurden. Auch zu Gunsten der Ökonomie der Befragungsdurchführung könnte ein Lösungsvorschlag darin liegen, auf die negativ formulierten Items zu verzichten. Zumindest sollten die Antworten der Kinder auf diese Items gesondert auf mögliche Antworttendenzen hin analysiert werden. Es bleibt zudem zu überlegen, inwieweit differenzielle Normen für Jungen und Mädchen bzw. für Kinder mit und ohne Förderbedarf notwendig sind. Weiterhin unterliegen die Ergebnisse Klasseneffekten, die bei der Bewertung berücksichtigt werden müssen.

Die hier dargelegten Schlüsse sollten vor dem Hintergrund verschiedener Einschränkungen gesehen werden. So führt die vorgenommene Modellspezifikation nicht zu gut annehmbaren Fit-Indices. Mögliche Grün-

de hierfür sind in Doppelladungen oder zu schwache Ladungen zu sehen. Zudem zeigte sich eine hohe Variabilität zwischen den Items, sodass zur Modellbeschreibung auf kongenerische Modelle (freie Faktorladungen) zurückgegriffen werden sollte. Detailliertere Itemanalysen sollten anhand von Modellen der Item Response Theorie erfolgen. Für das hier verfolgte Ziel sind diese Aspekte jedoch weniger bedeutsam und wurden entsprechend nicht differenziert geprüft.

In der vorliegenden Studie wurde eine breite Definition des SPF verwendet. Diese berücksichtigt mehrere schulische Entwicklungsproblematiken (Klassenwiederholung, attestierte Förderbedarfe im Umfang einer Teilleistungsstörung bzw. eines sonderpädagogischen Förderbedarfs). Dieses Vorgehen lässt keine Untersuchung differenzieller Effekte bestimmter Förderbedarfe zu, wird jedoch der Heterogenität der Gruppen von Kindern mit unterschiedlichen Förderbedarfen gerecht.

## Literatur

- Alansari, M., & Rubie-Davies, C. (2020). What about the tertiary climate? Reflecting on five decades of class climate research. *Learning Environments Research*, 23(1), 1–25. <https://doi.org/10.1007/s10984-019-09288-9>
- Aldrich, J. M., Fraser, B. J., Fozdar, F., Ala'i, K., Earnest, J. & Afari, E. (2016). Students' perceptions of school climate as determinants of wellbeing, resilience and identity. *Improving Schools*, 19(1), 5-26.
- Begert, T. & Müller, C. M. (2019). Zum Einfluss des Klassenzusammenhalts auf Unterrichtsstörungen. Eine netzwerkanalytische Längsschnittstudie auf der Sekundarstufe I. *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete*, 88(2), 100 –113. <https://doi.org/10.2378/vhn2019.art19d>
- Beld, M. H. M., Van den Heuvel, E. G., van der Helm, G. H. P., Kuiper, C. H. Z., de Swart, J. J. W., Roest, J. J. & Stams, G. J. J. M. (2019). The impact of classroom climate on students' perception of social exclusion in secondary special education. *Children and Youth Services Review*, 103, 127-134. <https://doi.org/10.1016/j.childyouth.2019.05.041>
- Blatchford, P., Bassett, P. & Brown, P. (2005). Teachers' and pupils' behavior in large and small classes: A systematic observation study of pupils aged 10 and 11 years. *Journal of Educational Psychology*, 97, 454-467.
- Blumenthal, Y. & Blumenthal, S. (2021). Zur Situation von Grundschülerinnen und Grundschulern mit sonderpädagogischem Förderbedarf im Bereich emotionale und soziale Entwicklung im inklusiven Unterricht. Longitudinale Betrachtung von Klassenklima, Lehrer-Schüler-Beziehung und sozialer Partizipation. *Zeitschrift für pädagogische Psychologie*. <https://doi.org/10.1024/1010-0652/a000323>
- Bogner, K. & Landrock, U. (2015). *Antworttendenzen in standardisierten Umfragen*. Mannheim, GESIS Leibniz Institut für Sozialwissenschaften (GESISsurvey Guidelines). [https://doi.org/10.15465/gesis-sg\\_016](https://doi.org/10.15465/gesis-sg_016)
- Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd Ed.). New York: Guilford Press.
- Campbell, D. T., & Kenny, D. A. (1999). *A primer on regression artifacts*. Guilford Press.
- Cheung, G. W. & Rensvold, R. B. (1998). Cross-Cultural Comparisons Using Non-Invariant Measurement Items. *Applied Behavioural Science Review*, 6, 93–110. [https://doi.org/10.1016/S1068-8595\(99\)80006-3](https://doi.org/10.1016/S1068-8595(99)80006-3)

- Crede, J., Wirthwein, L., Steinmayer, R. & Bergold, S. (2019). Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf im Bereich emotionale und soziale Entwicklung und ihre Peers im inklusiven Unterricht - Unterschiede in sozialer Partizipation, Schuleinstellung und schulischem Selbstkonzept. *Zeitschrift für Pädagogische Psychologie*, 33(3-4), 207-221.
- Crede, M. & Niehorster, S. (2012). Adjustment to College as Measured by the Student Adaptation to College Questionnaire: A Quantitative Review of its Structure and Relationships with Correlates and Consequences. *Educational Psychology Review*, 24, 133-165.
- Deegan, J. (1978). On the Occurrence of Standardized Regression Coefficients Greater Than One. *Educational and Psychological Measurement*, 38(4), 873-888. <https://doi.org/10.1177/001316447803800404>
- Eder, F. (1996). *Schul- und Klassenklima*. Innsbruck: Studienverlag.
- Eder, F. (2018). Schul- und Klassenklima. In D. H. Rost, J. R. Sparfeldt & S. R. Buch (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (S. 696 – 707). Weinheim: Psychologie Verlags Union.
- Fend, H. (1977). *Schulklima*. Weinheim: Beltz.
- Gentry, M., Gable, R. K., & Rizza, M. G. (2002). Students' perceptions of classroom activities: Are there grade-level and gender differences? *Journal of Educational Psychology*, 94(3), 539-544. <https://doi.org/10.1037/0022-0663.94.3.539>
- Grewe, N. (2003). *Aktive Gestaltung des Klassenklimas: Eine empirische Interventionsstudie*. Münster: Lit.
- Hascher, T. & Hagenauer, G. (2018). Die Bedeutung von Qualitätsfaktoren des Unterrichts und Lernemotionen für das Wohlbefinden in der Schule. In G. Hagenauer & T. Hascher (Hrsg.), *Emotionen und Emotionsregulation in Schule und Hochschule* (S. 103-119). Münster; New York: Waxmann.
- Hascher, T. (2017). Die Bedeutung von Wohlbefinden und Sozialklima für Inklusion. In B. Lütje-Klose, S. Miller, S. Schwab & B. Streese (Hrsg.), *Inklusion: Profile für die Schul- und Unterrichtsentwicklung in Deutschland, Österreich und der Schweiz* (S. 69 – 79). Münster: Waxmann.
- Henke, T., Bosse, S., Lambrecht, J., Jäntschi, C., Jaeuthe, J. & Spörer, N. (2017). Mit-tendrin oder nur dabei? Zum Zusammenhang zwischen sonderpädagogischem Förderbedarf und sozialer Partizipation von Grundschülerinnen und Grundschülern. *Zeitschrift für Pädagogische Psychologie*, 31, 111 – 123. <https://doi.org/10.1024/1010-0652/a00019>
- Hoffmann, L., Wilbert, J., Lehofer, M. & Schwab, S. (2020). Are we good friends? – Friendship preferences and the quantity and quality of mutual friendships. *European Journal of Special Needs Education*, 1-15. <https://doi.org/10.1080/08856257.2020.1769980>
- Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2. Aufl.). New York: Routledge.
- Huber, C. & Wilbert, J. (2012). Soziale Ausgrenzung von Schülern mit sonderpädagogischem Förderbedarf und niedrigen Schulleistungen im gemeinsamen Unterricht. *Empirische Sonderpädagogik*, 2, 147-165.
- Huber, C., Gebhardt, M. F. & Schwab, S. (2015). Lehrkraftfeedback oder Spaß beim Spiel? Eine Experimentalstudie zum Einfluss von Lehrkraftfeedback auf die soziale Akzeptanz bei Grundschulkindern. *Psychologie in Erziehung und Unterricht*, 62(1), 51-64.
- Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. & Rosseel, Y. (2020). *semTools: Useful tools for structural equation modeling*. R package version 0.5-3. Retrieved from <https://CRAN.R-project.org/package=semTools>
- Kenny, D. A. (1979). *Correlation and Causality*. John Wiley & Sons.

- Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford publications.
- Köb, S. & Janz, F. (2021). Teachers' subjective theories regarding social dynamics management and social participation of students with Intellectual Disabilities in inclusive classrooms. *International Journal of Inclusive Education*. <https://doi.org/10.1080/13603116.2021.1902002>
- Krawinkel, S., Südkamp, A., Lange, S. & Troster, H. (2017). Soziale Partizipation in inklusiven Grundschulklassen: Bedeutung von Klassen- und Lehrkraftmerkmalen. *Empirische Sonderpädagogik*, 9, 277 – 295.
- Mamas, C., Schaelli, G., Daly, A., Navarro, H. & Trisokka, L. (2019). Employing Social Network Analysis to Examine the Social Participation of Students Identified as Having Special Educational Needs and Disabilities. *International Journal of Disability, Development and Education*, 1–16. <https://doi.org/10.1080/1034912X.2019.1614153>
- Moosbrugger, H. & Kelava, A. (Hrsg.). (2012). *Testtheorie und Fragebogenkonstruktion* (2., akt. und überarb. Aufl.). Berlin, Heidelberg: Springer-Verlag.
- Morin, A. J. S., Marsh, H. W., Nagengast, B. & Scalas, L. F. (2014). Doubly Latent Multilevel Analyses of Classroom Climate: An Illustration, *The Journal of Experimental Education*, 82(2), 143-167. <https://doi.org/10.1080/00220973.2013.769412>
- R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available online at <https://www.R-project.org>.
- Rathmann, K., Herke, M., Hurrelmann, K. & Richter, M. (2018). Klassenklima, schulisches Wohlbefinden und Gesundheit von Schülerinnen und Schülern in Deutschland: Ergebnisse des Nationalen Bildungspanels (NEPS). *Gesundheitswesen*, 80(4), 332-341. <https://doi.org/10.1055/s-0043-121886>
- Rauer, W. & Schuck, K.-D. (2003). *Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern dritter und vierter Klassen* (FEES 3–4). Göttingen: Hogrefe.
- Roorda, D. L., Koomen, H. M. Y., Spilt, J. L. & Oort, F. J. (2011). The influence of affective teacher-student relationships on students' school engagement and achievement: a meta-analytic approach. *Review of Educational Research*, 81, 483-529.
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1–36.
- Rutkowski, L. & Svetina, D. (2014). Assessing the Hypothesis of Measurement Invariance in the Context of Large-Scale International Surveys. *Educational and Psychological Measurement*, 74, 31–57. <https://doi.org/10.1177/0013164413498257>
- Schurig, M., Glesemann, B., & Schröder, J. (2016). Dimensionen von Unterrichtsqualität: die Generalisierbarkeit von Schülerurteilen über Fächer und Zeit. In R. Strietholt, W. Bos, H. G. Holtappels, & N. McElvany (Hrsg.), *Jahrbuch der Schulentwicklung* (Band 19, S. 30–56). Weinheim: Beltz.
- Spörer, N., Maaz, K., Vock, M., Schröder-Lenzen, A., Luka, T., Bosse, S. et al. (2015). Lernen in der inklusiven Grundschule: Zusammenhänge zwischen fachlichen Kompetenzen, Sozialklima und Facetten des Selbstkonzepts. *Unterrichtswissenschaft*, 43, 22 – 35.
- Thapa A., Cohen J., Guffey S. & Higgins-D'Alessandro, A. (2013). A Review of School Climate Research. *Review of Educational Research*, 83(3), 357-385. <https://doi.org/10.3102/0034654313483907>
- Vogl, S. (2015). *Interviews mit Kindern führen*. Weinheim: Beltz.
- von Oertzen, T., Brandmaier, A. M. & Tsang, S. (2015). Structural Equation Modeling with  $\Omega$ nyx. *Structural Equation Modeling: A Multidisciplinary Journal*, 22(1), 148-161, <https://doi.org/10.1080/10705511.2014.935842>

Wilbert J., Urton, K., Krull, J., Kulawiak, P. R., Schwalbe, A. & Hennemann, T. (2020). Teachers' Accuracy in Estimating Social Inclusion of Students With and Without Special Educational Needs. *Frontiers in Education*, 5:598330. <https://doi.org/10.3389/educ.2020.598330>

### **AutorInnenhinweis**

*Korrespondenzadresse:*

Stefan Blumenthal

Stefan.blumenthal@uni-rostock.de

*Zum Manuskript existiert ein elektronisches Supplement mit detaillierteren Ergebnissen und dem Datensatz nebst R-Syntax. Das elektronische Supplement ist unter <https://www.doi.org/10.17605/OSF.IO/HBFSG> erreichbar.*

Erstmals eingereicht am 13.07.2021

Letzte Überarbeitung eingereicht am  
06.04.2022

Angenommen am 25.05.2022