Zitiervorschlag: Pham, G., Freunberger, R., Robitzsch, A., Itzlinger-Bruneforth, U., & Bruneforth, M. (2016). Reliabilität und Stabilität des Index der sozialen Benachteiligung und Kompositionseffekt der Schulen. Zeitschrift für Bildungsforschung, 6(3), 345–364. <https://doi.org/10.1007/s35834-016-0164-1>

Zur Verfügung gestellt auf PHIQ:

PHIQ-DOI: <https://doi.org/10.18747/PHSG-coll3/id/741>

Original-DOI: <https://doi.org/10.1007/s35834-016-0164-1>

Dokumentart: Journal article

Version: accepted version

Copyright-Hinweis: "This is a post-peer-review, pre-copyedit version of an article published in Zeitschrift für Bildungsforschung. The final authenticated version is available online at: <https://doi.org/10.1007/s35834-016-0164-1>."

Lizenz: Alle Rechte vorbehalten

Zusammenfassung

Zum Ausgleich erschwelter Bedingungen an Schulen mit ungünstiger sozialer Komposition wird vielfach eine Kompensation durch sozialindexbasierte Finanzierung diskutiert. Ein möglicher Verteilungsschlüssel für die Finanzierung ist der Index der sozialen Benachteiligung (ISB). Der ISB erfasst die soziale und ethnische Zusammensetzung der österreichischen Schulen und wurde im Rahmen der Bildungsstandardüberprüfung (BIST-Ü) 2013 erstmals in den Ergebnisberichten rückgemeldet. Die Analyse der Reliabilität dieses Index sowie der Stabilität über zwei Erhebungszeitpunkte hinweg ist Gegenstand dieses Artikels. Des Weiteren soll der Effekt der Zusammensetzung der Schülerschaft in Schulen auf die Schülerleistung anhand der BIST-Ü-Daten aus den Jahren 2012 und 2013 betrachtet werden. Die Ergebnisse – basierend auf deskriptiven Statistiken und Intraklassenkorrelationen – zeigen, dass der Index eine zufriedenstellende Reliabilität aufweist. Allerdings stellt dieser aggregierte Index erwartungsgemäß für Kleinschulen ein weniger reliables Maß dar als für Schulen mit höheren Schülerzahlen auf der jeweiligen Schulstufe. Die Ergebnisse aus einem Dreiebenenmodell sowie manifeste Veränderungsmaße zeigen weiters, dass der ISB ein stabiles Maß über die zwei Erhebungszeitpunkte ist. Mithilfe eines Mehrebenenmodells mit latenten Kovariaten bestätigt sich der Effekt des ISB auf die Schülerleistung sowohl auf der Primarstufe als auch auf der Sekundarstufe I. Auf der Sekundarstufe I ist dieser auch nach Kontrolle der Schulform signifikant.

Schlüsselwörter: Sozialindexbasierte Finanzierung, Index der sozialen Benachteiligung, Reliabilität, Stabilität, Kompositionseffekt, Mehrebenenanalyse

Abstract

School funding formulas have been much discussed to compensate the needs of socially disadvantaged schools. The index of social disadvantages (“Index der sozialen Benachteiligung”, ISB), a measure of the social and ethnical composition in Austrian schools, was suggested as a key variable for this purpose. It was reported in the context of the national standards-based assessment tests in 2013. The main topic of the presented article is to provide insights into the reliability and stability (over two consecutive measurement occasions) of the proposed index. Furthermore, we evaluated the school composition effect, indicating the relationship between school composition and performance. For all analyses data from the national standards-based assessment tests from 2012 and 2013 were used. The results – based on descriptive statistics and intraclass correlation coefficients – show that the school index is a sufficiently reliable measure. As expected however, the index proves to be less reliable for small schools than for schools with a higher number of students. In addition, as revealed by a multilevel analysis and conventional change measures, the index proofs to be stable over the two measurement occasions. Finally, the results of a multilevel latent covariate model show evidence for a composition effect, which holds for both primary and lower secondary schools. Also, the composition effect is evident even after controlling for school track.

Key words: school funding formula, index of social disadvantages, reliability, stability, composition effect, multilevel analysis

1 Einleitung und Fragestellung

Das Prinzip der Chancengerechtigkeit ist ausdrücklich als Ziel in der österreichischen Bundesverfassung verankert und bedeutet, dass der Bildungserfolg der Lernenden nicht von Herkunft, sozialer Lage und finanziellem Hintergrund der Familie abhängen soll (Art. 14 B-VG i. d. F. BGBl. I Nr. 31/2005). Der Nationale Bildungsbericht 2012 hat jedoch gezeigt, dass dieses Ziel nicht erreicht wurde: Schullaufbahn und Bildungserfolg hängen klar von der sozialen Herkunft und vom Migrationshintergrund ab (Bruneforth, Weber & Bacher, 2012). Dies gilt nicht nur auf der Ebene individueller Schülermerkmale, sondern zusätzlich auch auf der Ebene der Zusammensetzung von Schulklassen und Schulen. Die letztgenannten Effekte werden in der Bildungsforschung *Kompositionseffekte* genannt (für einen Überblick der einschlägigen internationalen Forschung siehe Dumont, Neumann, Maaz und Trautwein, 2013; mit besonderem Bezug zu Österreich siehe Biedermann, Weber, Herzog-Punzenberger und Nagel, 2016). In der Praxis bedeutet dies, dass Schulen infolge unterschiedlicher sozialer Zusammensetzung ihrer Schülerschaft unterschiedliche Startchancen für erfolgreiche Lehr- und Lernprozesse haben. Eine mögliche bildungspolitische Konsequenz dieses Sachverhalts ungleicher Ausgangsbedingungen ist die Kompensation durch eine zusätzliche finanzielle Unterstützung; zur aktuellen Diskussion zu diesem Themenbereich in Österreich siehe Kuscheij und Schönpflug (2014) sowie Bacher (2016). Zur Frage, ob und in welcher Weise die soziale Benachteiligung von Schulen durch ungünstige Kompositionseffekte ausgeglichen werden soll, gibt es eine vielfältige internationale Diskussion (siehe Fazekas, 2012; Levačić, 2008). Für Österreich haben Bacher, Altrichter und Nagy (2010) ein Modell einer sozialindexbasierten Finanzierung mit einem Sozialindex als Kern vorgeschlagen, darauf aufbauend wurde der Index von Bruneforth et al. (2012) weiterentwickelt.

In diesem Artikel geht es *nicht* darum, das Thema „Chancengerechtigkeit“ theoretisch oder historisch zu vertiefen. Ebenso wenig soll hier ein weiterer Sozialindex vorgeschlagen werden. Vielmehr möchten wir den von Bruneforth et al. (2012) vorgestellten Sozialindex daraufhin überprüfen, ob er zentralen psychometrischen Gütekriterien in Bezug auf den genannten Zweck genügt. Dieser *Index der sozialen Benachteiligung* (ISB) reflektiert die ungünstige soziale und ethnische Zusammensetzung der Schülerschaft von Schulen und soll den zusätzlichen Mittelbedarf benachteiligter Schulen abbilden. 2013 wurde der ISB erstmals im Rahmen der österreichischen Bildungsstandardüberprüfung (BIST-Ü) systematisch auf System- und Schulebene rückgemeldet (Schreiner & Breit, 2014a, S. 37–39; 2014b, S. 65–67). Die Berechnungsweise und die Datengrundlage dieses Index werden in den Abschnitten 2.1 und 2.2 näher erläutert.

Die Hauptfragestellung, nämlich inwieweit sich der ISB auf Grundlage der BIST-Ü-Daten aus psychometrischer Sicht als Basis für ein Finanzmodell im Schulbereich eignen würde, bezieht sich auf die von Bacher (2015) genannten Anforderungen an einen solchen Index einschließlich Zuverlässigkeit, Objektivität, Gültigkeit, Sensitivität und Unverzerrtheit. In diesem Beitrag beschränken wir uns auf zwei Aspekte. Der erste Aspekt bezieht sich auf die Gültigkeit des Index als einen repräsentativen Wert der Schule, welcher zumindest über folgende Eigenschaften verfügt: Unterscheidbarkeit zwischen Schulen, Reliabilität und Stabilität. Der zweite Aspekt bezieht sich auf die Gültigkeit des Index als Basis für ein Finanzmodell im Pflichtschulbereich, welches den Zusatzbedarf der Schulen messen soll: Mit der Feststellung einer unterschiedlich ungünstigen Ausgangslage bezüglich der Schülerzusammensetzung ist implizit die Hypothese verbunden, dass dies negative Auswirkungen

auf die Schülerleistung hat. Diese Hypothese soll empirisch überprüft werden. Damit verbunden lauten die Forschungsfragen unserer Studie wie folgt:

Forschungsfrage 1: Wie ungleich sind individuelle Schülermerkmale innerhalb von Schulen in Österreich verteilt (Heterogenität der Schülerschaft der Schulen)? Auf Basis bisheriger Befunde (z. B. Bacher, 2015; Bruneforth & Lassnigg, 2012, S. 36–39; Breit & Schreiner, 2006, S. 196) erwarten wir deutliche Differenzen hinsichtlich des ISB im Vergleich der Schulstandorte, aber auch zwischen Schulformen einerseits und den Bundesländern andererseits.

Forschungsfrage 2: Wie reliabel kann der ISB als aggregierter Wert der Schule gemessen werden? Da die Reliabilität eines Gruppenmittelwerts von der Gruppengröße abhängt (Raudenbush & Bryk, 2002, S. 40), gehen wir davon aus, dass der ISB der Schulen auf Basis der Daten von wenigen Schülerinnen und Schülern nicht besonders reliabel ist und die Reliabilität des ISB von großen Schulen ausreichend hoch ausfällt.

Forschungsfrage 3: Ist der ISB ein stabiles Maß über mehrere Schülerkohorten (äquivalent zu Untersuchungsjahre) hinweg oder hängt er von der untersuchten Kohorte ab? Weil es in den Untersuchungsjahren 2012 und 2013 keine auffällige Populationsänderung in Österreich gegeben hat, wird Ersteres erwartet.

Forschungsfrage 4: Inwieweit ist die Annahme eines negativen Zusammenhangs des ISB der Schule mit den Schülerleistungen berechtigt? Erreichen Schüler/innen in Schulen mit höherem ISB geringere Leistungen als solche in Schulen mit niedrigerem ISB (ISB-Effekt, *Frage 4a*)? Auf Basis bisheriger Ergebnisse (z. B. Bruneforth et al., 2012) erwarten wir, dass diese Frage bejaht wird. Ferner wollen wir überprüfen, ob der ISB-Effekt der Schulen über den durch individuelle Schülermerkmale bedingten Effekt hinausgeht (Kompositionseffekt des ISB, *Frage 4b*). Angesichts der aktuellen Befunde zu sozialen und ethnischen Kompositionseffekten der Schulen (für Österreich siehe Biedermann et al., 2016) lässt sich vermuten, dass diese Frage auch mit Ja beantwortet wird. Mit Blick auf die für Deutschland mehrfach bestätigte starke Konfundierung zwischen den sozialen und ethnischen Kompositionseffekten und den Institutionseffekten (institutionelle Unterschiede zwischen Schulformen, vgl. Gröhlich, Guill, Scharenberg & Bos, 2010; Maaz, Trautwein, Lüdtke & Baumert, 2008; Baumert, Stanat & Watermann, 2006) kann man jedoch hinterfragen, ob der angenommene Zusammenhang zwischen dem ISB und der Schulleistung auf der Sekundarstufe I nicht vollständig durch die schulformbezogenen Unterschiede erklärt wird. Aus diesem Grund möchten wir weiters untersuchen, ob der Effekt und der Kompositionseffekt des ISB auf die Schülerleistung auch nach Kontrolle der Schulform existieren (*Frage 4c*). Im Einklang mit bisherigen Befunden nehmen wir an, dass sich der angenommene Effekt des ISB nach Kontrolle der Schulform stark verringert und nicht mehr bedeutsam ist.

2 Methodik

2.1 Berechnungsweise des Index der sozialen Benachteiligung

Der ISB setzt sich nach dem Vorschlag von Bruneforth et al. (2012) aus vier Indikatoren zusammen: Bildung der Eltern, Beruf der Eltern (HISEI, *Highest International Socio-Economic Index of occupational status*, Ganzeboom, De Graaf & Treiman, 1992), Migrationsstatus und Muttersprache des Kinds. Die Auswahl dieser

Indikatoren erfolgte nach empirischen Analysen der Autoren anhand von Daten der Ausgangsmessungen für die BIST-Ü und aufgrund ihrer bedeutsamen Einflüsse auf den Schul- und Bildungserfolg in Österreich sowohl auf individueller Schülerebene als auch auf Klassen- und Schulebene (vgl. Bacher, 2016). Für jede Schülerin/jeden Schüler werden vier Variablen (d) jeweils mit Wert 0 (Merkmal trifft nicht zu) oder 1 (Merkmal trifft zu) berechnet: Schüler/in hat Migrationshintergrund¹ (Migrationshintergrund, d_{1ij}), Schüler/in hat andere Muttersprache als Deutsch (Muttersprache, d_{2ij}), Eltern von Schüler/in haben maximal Pflichtschulabschluss (Elternausbildung, d_{3ij}), Schüler/in stammt aus unterem Quintil² des HISEI (HISEI, d_{4ij}). Der Index B_{ij} auf Individualebene berechnet sich für eine Schülerin/einen Schüler i in Schule j durch den Mittelwert dieser vier Variablen als Basis, mit 100 multipliziert und mit einer „Basiszahl“ (bezieht sich auf den Basisbedarf der Schulen, vgl. Bacher et al., 2010) von 100 addiert. Für jede Schülerin/jeden Schüler können sich somit Werte von 100, 125, 150, 175 und 200 ergeben. Darauf basierend wird der Schulindex $\bar{B}_j$ als Mittelwert aller individuellen Schülerwerte B_{ij} berechnet. Aufgrund ihrer Einfachheit, Interpretierbarkeit und Kommunizierbarkeit wurde diese Berechnungsweise anstatt komplexerer Alternativen wie beispielsweise Faktorenanalyse oder Rasch-Skalierung vorgeschlagen (Bacher, 2015).

2.2 Datengrundlage

Als Basis für die Analysen standen die Daten aus der Bildungsstandardüberprüfung (BIST-Ü) der Jahre 2012 und 2013 für die Fächer Mathematik auf der vierten (M4 2013) und der achten Schulstufe (M8 2012) sowie Englisch-Lesen auf der achten Schulstufe (E8 2013) zur Verfügung. Die Zielpopulation der BIST-Ü umfasst alle Schüler/innen der jeweiligen Schülerkohorte von allen öffentlichen sowie privaten Volksschulen (VS) mit Öffentlichkeitsrecht auf der vierten Schulstufe bzw. Schüler/innen der achten Schulstufe an allen öffentlichen Schulen sowie an privaten Schulen mit gesetzlich geregelten Schularthbezeichnungen (Volksschuloberstufe, Hauptschule und Neue Mittelschule – allgemeinbildende Pflichtschule, APS, und allgemeinbildende höhere Schule, AHS) und Öffentlichkeitsrecht. An der BIST-Ü M4 2013 nahmen insgesamt 73 655 Viertklässler/innen aus 3048 VS teil. Auf der Sekundarstufe I (Sek. I) nahmen 76 728 Achtklässler/innen aus 1410 Schulen an der BIST-Ü E8 2013 und 79 678 Achtklässler/innen aus 1415 Schulen an der BIST-Ü M8 2012 teil. Von der Überprüfung ausgenommen waren alle außerordentlichen Schüler/innen sowie Schüler/innen mit sonderpädagogischem Förderbedarf, die im Testfach nach dem Lehrplan der allgemeinen Sonderschule oder nach dem Lehrplan einer niedrigeren Schulstufe unterrichtet wurden. Schüler/innen mit Körper- oder Sinnesbehinderung nahmen nicht an der Überprüfung teil, wenn sie selbst mit allenfalls im Unterricht zur Verfügung stehenden Unterrichts- oder Hilfsmitteln unter den standardisierten Testbedingungen die gestellten

¹ Die Definition des Migrationshintergrunds erfolgt in Anlehnung an jene der OECD, welche als Kriterium das Geburtsland der Eltern heranzieht. Ein Kind gilt demnach als Schüler/in mit Migrationshintergrund, wenn beide Elternteile im Ausland geboren wurden. Eine Ausnahme von dieser Regel bilden im Bericht Schüler/innen, deren Eltern (ein Elternteil oder beide Elternteile) in Deutschland geboren sind – sie werden aufgrund der gleichen Sprache nicht zur Gruppe der Schüler/innen mit Migrationshintergrund gezählt.

² 20% aller Werte sind kleiner als dieser Wert.

Aufgaben aller Voraussicht nach nicht hätten lösen können. Im Erhebungsjahr 2013 betrug der durchschnittliche Anteil ausgeschlossener Schüler/innen pro Schule 3.7% auf der vierten Schulstufe und 3.3% auf der achten Schulstufe. Eine Übersicht der Teilnehmerstatistik der drei BIST-Ü in den Jahren 2012 und 2013 findet sich in Tabelle 1.

[Tabelle 1 hier einfügen]

Zwischen Schulen variierte dieser Schüleranteil stark: zwischen 0% und 67% auf der Volksschulstufe sowie zwischen 0% und 50% auf der Sek. I. Der Ausschluss dieser Schüler/innen bei der BIST-Ü führte dazu, dass sie bei der Berechnung des ISB nicht berücksichtigt wurden. Insgesamt betraf dies pro Studie zwischen 2697 und 3541 Schüler/innen.

Die Schülerleistung wurde mithilfe standardisierter Tests im Rahmen der BIST-Ü erfasst. Die Leistungsdaten im Fach Mathematik wurden nach dem Rasch-Modell (Rasch, 1960), im Fach Englisch nach dem Partial-Credit-Modell (Masters, 1982) skaliert. Für die weiteren Analysen wurden alle Testwerte z -standardisiert (d. h. $M = 0$ und $SD = 1$), damit die Ergebnisse verschiedener Schulstufen und Fächer sich auf der gleichen Metrik befinden.

Die Daten für die Bildung des ISB stammen hauptsächlich aus dem Schülerfragebogen der jeweiligen Studie, lediglich die Angaben zur Elternausbildung in der BIST-Ü M4 2013 stammen aus dem Elternfragebogen. Fehlende Schüler- und Elternangaben (Missings) sind vorhanden. In Tabelle 2 werden die Missinganteile der vier Indikatoren des ISB dargestellt. Hier entsprechen 100% der Anzahl aller getesteten Schüler/innen der jeweiligen Überprüfung.

[Tabelle 2 hier einfügen]

Aus Tabelle 2 wird ersichtlich, dass die Missinganteile der vier Indikatoren zum Teil hoch sind. Am höchsten ist der Missinganteil bei den Angaben zur Elternausbildung, auf der achten Schulstufe sind dies sogar mehr als zehn Prozent. Da fehlende Werte zur verzerrten Berechnung des ISB führen können, wurden diese imputiert. Als Datenbasis für die Analysen in diesem Beitrag verwenden wir die 50 imputierten Datensätze aus dem Prozess des fairen Vergleichs der jeweiligen Überprüfungen (siehe George & Robitzsch, 2015; Pham & Robitzsch, 2015). Dabei wurde angenommen, dass die Daten zufällig fehlten (*missing at random*, Rubin, 1976). Das bedeutet, dass das Auftreten von fehlenden Werten einer Variable durch die Ausprägung anderer beobachteter Variablen im Imputationsmodell, nicht jedoch durch die Ausprägung der Variable selbst (nach Berücksichtigung der beobachteten Variablen) bedingt ist. Zur Datenimputation wurde ein Random-Intercept-Modell unter Berücksichtigung der Mehrebenenstruktur der Daten (Schüler/innen geschachtelt innerhalb von Schulen) mit Hilfe des R-Pakets `pan` (Zhao & Schafer, 2013) durchgeführt. Als Kovariaten im Modell wurden die Kontextvariablen der Schule (u. a. Gemeindegröße, Entfernung zur nächstgelegenen AHS, Art des Schulerhalters, Schulgröße³) sowie die Zusammensetzung der Schülerschaft der Schule (u. a. Mädchenanteil, Anteil der Schüler/innen mit Migrationshintergrund, mittlerer sozioökonomischer Status) einbezogen. Durch die

³ Der Begriff „Schulgröße“ bezieht sich in unserem Beitrag auf die getestete Schülerzahl der betroffenen Schulstufe der einzelnen Schulen. Es entspricht in der VS jeweils ca. einem Viertel der tatsächlichen Schulgröße und auf der Sek. I ca. einem Viertel der Schülerzahl der Schulstufen 5 bis 8.

multiplen Imputationen kann die Unsicherheit bei der Ersetzung der Missings berücksichtigt werden (vgl. Lüdtke & Robitzsch, 2010).

2.3 Statistische Analysen

Für die Analysen kommt das von Rubin (1987) beschriebene Verfahren zum Einsatz; jede Berechnung wird zuerst mit jedem der 50 imputierten Datensätzen durchgeführt, dann werden die Ergebnisse gepoolt und berichtet. Die Analysen der ersten drei Forschungsfragen werden in der Software R (R Core Team, 2015) durchgeführt. In R werden die R-Pakete `psych` (Revelle, 2014), `semTools` (Pornprasertmanit, Miller, Schoemann & Rosseel, 2013), `Hmisc` (Harrell, 2015), `lme4` (Bates, Maechler, Bolker & Walker, 2015) und `BIFIEsurvey` (BIFIE, 2015) verwendet. Bis zum Zeitpunkt der Analysen der Forschungsfragen 4a, 4b und 4c ist uns kein entsprechendes R-Paket bekannt, daher verwenden wir hierfür die Software Mplus (Muthén & Muthén, 2012).

Um die erste Fragestellung zu beantworten, wird neben den deskriptiven Statistiken (Mittelwert M , Standardabweichung SD und Quantile) auch die ICC (*Intraclass Correlation Coefficient*) als Überprüfungsmaß vorgeschlagen (Raudenbush & Bryk, 2002, S. 36). Die ICC gibt den Anteil der Gesamtvarianz an, der durch Unterschiede zwischen Schulen erklärt wird, und zeigt, wie stark die Schulen sich voneinander unterscheiden.

Die Reliabilität des Schulmittelwerts ICC(2) berechnet sich nach Anwendung der Spearman-Brown-Formel auf Basis der ICC (siehe Snijders & Bosker, 2012, S. 26). Damit wird die Varianz innerhalb der Schulen mit der jeweiligen Schulgröße n_j gewichtet. Man erhält damit eine Aussage darüber, inwieweit sich die Reliabilität des Schulmittelwerts unter Annahme einer bestimmten Schulgröße verändern würde. Um eine „durchschnittliche“ Reliabilität für die gesamte Stichprobe zu ermitteln, kann n_j durch die mittlere Schulgröße $\bar{n}$ ersetzt werden.

Um zu beantworten, ob der Schulindex ein stabiles Maß der Schulen über mehrere Schülerkohorten hinweg ist (*Forschungsfrage 3*), werden die Daten der Sek. I verwendet (siehe Abschnitt 2.1). Dafür können verschiedene Maße berechnet werden: Stabilitätskoeffizienten, absolute Veränderungsmaße (Differenzwert) und relatives Übereinstimmungsmaß (Korrelation). Da auf der Primarstufe bis zum Zeitpunkt der Erstellung dieses Beitrags lediglich die Ergebnisse einer BIST-Ü im Jahr 2013 zur Verfügung standen, kann diese Forschungsfrage nur anhand der Daten der Sekundarstufe I analysiert werden.

Um die Stabilitätskoeffizienten zu berechnen, wird ein Dreiebenenmodell (Snijders & Bosker, 2012, S. 67–68) mit Schülerinnen und Schülern innerhalb jeder Schülerkohorte (Achtklässler/innen eines Erhebungsjahrs) einer Schule als Ebene 1, Schülerkohorten (Achtklässler/innen verschiedener Erhebungsjahre) innerhalb jeder Schule als Ebene 2 und die Schulen als Ebene 3 definiert, mit:

$$B_{ikj} = \mu_{00} + \mu_{01} \cdot z_j + u_{kj} + \varepsilon_{ikj} = \mu_{0j} + u_{0j} + w_{kj} + \varepsilon_{ikj} \quad (1)$$
$$\text{Var}(u_{0j}) = \tau_0^2, \text{Var}(w_{kj}) = \tau_r^2, \text{Var}(\varepsilon_{ikj}) = \delta_i^2$$

B_{ikj} ist das Individualmerkmal von Schüler/in i in Kohorte k der Schule j . Der Effekt u_{kj} von Schülerkohorte k der Schule j vom (schulformspezifischen) Gesamtmittelwert wird durch den kohortenübergreifenden Schuleffekt u_{0j} über mehrere Schülerkohorten hinweg und die Abweichung w_{kj} der Schülerkohorte k (kohortenspezifischer

Effekt) vorhergesagt. Dadurch kann der Anteil der kohortenübergreifenden Varianz (τ_0^2) an der Gesamtvarianz zwischen Schulen ($\tau_0^2 + \tau_r^2$) als Stabilitätskoeffizient $v = \tau_0^2 / (\tau_0^2 + \tau_r^2)$ berechnet werden.

Darüber hinaus werden absolute und relative Veränderungsmaße des manifesten Schulindex in Betracht gezogen. Als absolutes Veränderungsmaß wird die Differenz der Indizes von Schulen zwischen zwei verschiedenen Erhebungsjahren T1 (2012) und T2 (2013) definiert. Dieses Maß gibt an, wie stark sich der ISB von Jahr zu Jahr ändert. Um zu überprüfen, ob die Schulindizes von T1 zu T2 im Durchschnitt variieren, wird ein *t*-Test mit abhängigen Stichproben durchgeführt. Schließlich wird die Korrelation der Schulindizes zwischen zwei Erhebungsjahren berechnet, was sich als ein relatives Stabilitätsmaß deuten lässt.

Die Forschungsfragen 4a und 4b zielen darauf ab zu klären, ob Unterschiede zwischen Schulen bezüglich des ISB etwaige Leistungsunterschiede vorhersagen und ob dieser Effekt stärker ausfällt als der durch individuelle Schülermerkmale der Schulen bedingte Effekt. Dafür wird ein *multilevel-latent-covariate*-Modell (Lüdtke, Marsh, Robitzsch, Trautwein, Asparouhov & Muthén, 2008) definiert. Dieses Modell berücksichtigt die Mehrebenenstruktur der Daten (Schüler/innen geschachtelt innerhalb von Schulen) sowie die Reliabilität des Prädiktors (ISB) auf Schulebene (vgl. Abschnitt 3.3) und kann mit der Software Mplus (Muthén & Muthén, 2012) berechnet werden. Dabei ist die Schülerleistung die abhängige Variable (*Y*), ISB die unabhängige Variable (*B*); es wird kein random-Slope modelliert. Die Modellgleichung lautet:

$$Y_{ij} = \mu + \beta_w B_{w,ij} + \beta_b B_{b,j} + \varepsilon_{ij} + \varepsilon_j \quad (2)$$

Die individuelle Schülerleistung Y_{ij} von Schüler/in i in Schule j wird sowohl durch die individuellen Schülermerkmale als auch durch den latenten Schulmittelwert (Schulwert nach der Reliabilitätskorrektur) vorhergesagt. Die Größe $B_{b,j}$ ist die Differenz zwischen dem latenten Schulmittelwert und dem Gesamtmittelwert (μ , Intercept) des ISB. $B_{w,ij}$ ist die Abweichung der individuellen Indizes vom latenten Schulwert (groupmean-zentriert). Der within-Regressionskoeffizient β_w beschreibt somit den Effekt der individuellen Schülermerkmale hinsichtlich des ISB innerhalb der Schulen (within-Ebene). Der between-Regressionskoeffizient β_b beschreibt den zu untersuchenden Effekt des ISB auf Schulebene (between-Ebene, Frage 4a). Der Kompositionseffekt (Frage 4b) ist dann gegeben durch $\beta_c = \beta_b - \beta_w$. Dieser tritt auf, wenn der between-Effekt stärker als der within-Effekt ist (Lüdtke et al., 2008).

Zur Beantwortung der Forschungsfrage 4c nach dem Effekt und dem Kompositionseffekt des ISB auf der Sek. I auf die Schülerleistung nach Kontrolle der Schulform wird zusätzlich die Schulform (AHS) als Kovariate auf between-Ebene im Modell spezifiziert.

3 Ergebnisse

3.1 Deskriptive Statistik des Schulindex

Eine Übersicht über die deskriptive Statistik des ISB für die Volksschulen (VS) in Österreich und für die Bundesländer findet sich in Tabelle 3, für die Schulen der Sekundarstufe I (Sek. I) in Tabelle 4.

[Tabelle 3 hier einfügen]

Für 3048 VS in Österreich im Jahr 2013 beträgt der durchschnittliche Schulindex $M = 112.4$ ($SD = 11.1$). Der Mittelwert der Bundesländer variiert von 108.5 (Kärnten) bis 128.2 (Wien). Wien ist das Bundesland mit dem mit Abstand höchsten durchschnittlichen ISB im Bereich der VS. Ca. 75% der Volksschulen österreichweit sowie in sechs Bundesländern haben einen Index bis zu 115 (Österreich: 75%-Quantil = 115.5, in sechs Bundesländern: 75%-Quantil < 115). Zusätzlich haben weniger als 5% der VS in sieben Bundesländern einen Schulindex größer als 135 (95%-Quantil < 135), ausgenommen Oberösterreich (95%-Quantil = 136.9) und Wien (75%-Quantil = 142.7).

[Tabelle 4 hier einfügen]

Auf der Sek. I ist der Mittelwert des Schulindex von 1410 Schulen im Jahr 2013 $M = 117.2$ ($SD = 13.5$) höher als auf der Primarstufe. Das 75%-Quantil ist in allen Bundesländern größer als 115, in nur vier Bundesländern (Burgenland, Kärnten, Niederösterreich und Tirol) ist das 95%-Quantil kleiner als 135. Österreichweit haben ca. 5% der Schulen einen ISB größer als 149.6. Die Statistiken unterscheiden sich jedoch deutlich zwischen den allgemeinbildenden Pflichtschulen (APS, $M = 118.9$, $SD = 13.8$) und allgemeinbildenden höheren Schulen (AHS, $M = 110.1$, $SD = 9.3$). 95% der AHS haben einen Schulindex kleiner als 129.8, während 25% der APS über einen Index höher als 122.7 verfügen. Ca. 5% der APS haben einen Index größer als 151.9, in Wien sogar höher als 161.1.

Die Korrelationen von den vier Indikatoren des Index miteinander und mit dem ISB werden in Tabelle 5 berichtet. Alle vier Indikatoren korrelieren auf beiden Ebenen paarweise positiv miteinander und mit dem ISB, wobei die Korrelationen auf Schulebene stärker ausgeprägt sind. Da die Schüler/innen mit Migrationshintergrund häufig auch eine andere Muttersprache als Deutsch haben, korrelieren die ersten beiden Indikatoren sehr hoch miteinander, insbesondere auf der Schulebene ($r = .93$ auf der Primarstufe, $r = .97$ auf der Sek. I) und folglich auch mit dem ISB. Mit den letzten Indikatoren korreliert der ISB in beiden Schulstufen und auf beiden Ebenen etwas schwächer.

[Tabelle 5 hier einfügen]

3.2 Heterogenität der Schülerschaft und Unterschiede zwischen Schulen und Schulformen hinsichtlich des ISB

Die deskriptiven Statistiken des ISB (siehe Tabelle 3 und Tabelle 4) zeigen, dass es Unterschiede zwischen Schulen bezüglich des ISB gibt, sowohl in den VS als auch auf der Sek. I und auch innerhalb jedes Bundeslands (VS in Kärnten und Sek.-I-Schulen im Burgenland sind am homogensten bzgl. des ISB mit $SD = 5.9$ bzw. $SD = 6.3$). Während ca. 25% der Schulen auf beiden Schulstufen einen geringen Index (kleiner als 105.3 bzw. kleiner als 108.4) haben, weisen andere 25% der Schulen einen Index höher als 115.5 (in den VS) bzw. 120.7 (auf der Sek. I) auf. Die 5% Schulen mit dem geringsten und dem höchsten Wert des ISB unterscheiden sich hinsichtlich der Zusammensetzung ihrer Schülerschaft deutlich voneinander.

Die ICC des ISB der österreichischen VS liegt bei .18, die der Schulen der Sek. I bei .26. Unter Berücksichtigung des Umstands, dass die durchschnittliche ICC hinsichtlich der Unterschiede zwischen Schulen in der Schülerleistung zwischen 0.08 und 0.15 beträgt (Teddle, Reynolds & Sammons, 2000, S. 67), gibt es also

bedeutsame Unterschiede zwischen österreichischen Schulen in Bezug auf den ISB. Auf der Sek. I ist die ICC insgesamt und auch nach Bundesländern höher als auf der Primarstufe. Betrachtet man die Werte der ICC in Tabelle 3 und Tabelle 4 pro Bundesland, wird jedoch ersichtlich, dass diese sehr heterogen sind. Im Burgenland sieht man z. B. eine ICC der VS von .05, was bedeutet, dass 5% der Variation im ISB auf die Variation zwischen den VS rückführbar sind. In Wien dagegen beträgt der Anteil der Variation zwischen den VS 28%. Auf beiden Schulstufen ist die ICC im Burgenland, in Kärnten, Tirol, Vorarlberg und in Niederösterreich gering. In diesen Bundesländern fällt die Heterogenität der Schülerschaft innerhalb jeder Schule relativ groß aus, zwischen Schulen ist diese gering.

Auf der Sek. I besitzen die APS eine höhere ICC sowie eine größere Streuung ($ICC = .24$, $SD = 13.8$) als die AHS ($ICC = .19$, $SD = 9.3$). Ebenfalls unterscheiden sich die Quantile der APS deutlich stärker als die der AHS. Dies bedeutet, dass die AHS homogener hinsichtlich der sozialen Herkunft ihrer Schülerschaft sind als die APS.

3.3 Die Reliabilität des Schulindex

Zur Beantwortung der Frage, inwieweit der ISB ein reliables Maß für die Schulen darstellt (Forschungsfrage 2), wurde die ICC(2) berechnet. Die Ergebnisse der VS sind in Tabelle 6 dargestellt.

[Tabelle 6 hier einfügen]

Die Reliabilität ICC(2) des ISB der VS bei einer durchschnittlichen Schulgröße $\bar{n} = 24$ beträgt .84. Da die Durchschnittsschulgröße der Bundesländer (kleinste mittlere Schulgröße mit $\bar{n} = 24$ im Burgenland, größte Schulgröße mit $\bar{n} = 53$ in Wien) stark streut, unterscheidet sich die Reliabilität des ISB der VS stark nach Bundesland (kleinste ICC(2) = .41 im Burgenland, höchste ICC(2) = .95 in Wien). Darüber hinaus spielt die Schulgröße wie erwartet eine bedeutsame Rolle für die reliable Messung des ISB jeder einzelnen Schule. Tabelle 6 vergleicht die Reliabilität des ISB einer VS durchschnittlicher Größe mit der Reliabilität bei einer Schulgröße von 20 bzw. 50. Für größere Schulen ($n_j = 50$) ist die Reliabilität des Index mit einem Wert von .92 zufriedenstellend und dies ist auch bei etwas kleineren Schulen ($n_j = 20$) mit .82 der Fall.

[Tabelle 7 hier einfügen]

Auf der Sek. I haben Schulen im Vergleich zur VS eine größere Schulgröße ($\bar{n} = 54$, s. Tabelle 7). Die Reliabilität des ISB ($ICC(2) = .95$) ist dementsprechend höher als bei den VS. Das Gleiche gilt für die einzelnen Bundesländer sowie für die APS und AHS. Selbst die geringste ICC(2) = .77 im Burgenland fällt hinreichend aus.

3.4 Die Stabilität des Schulindex auf der Sekundarstufe I

Auf der Sek. I sind bisher Daten der Bildungsstandardüberprüfung (BIST-Ü) von zwei beinahe vollständigen Schülerkohorten von 1396 Schulen⁴ vorhanden. Darauf basierend konnten wir die Stabilität des ISB auf Schulebene untersuchen.

⁴ Manche Schulen wurden vor der BIST-Ü E8 2013 aufgelöst bzw. zusammengelegt, andere Schulen sind nach der BIST-Ü M8 2012 neu gegründet worden.

Der Gesamtmittelwert des ISB der APS über zwei Kohorten hinweg betrug 119.0 Punkte, der der AHS war um 9.2 Punkte geringer. Die Ergebnisse des Dreiebenenmodells zeigen eine hohe Stabilität des ISB über Kohorten hinweg. Die Variation des ISB zwischen Messzeitpunkten auf Schulebene ist sehr klein, was sich in einem hohen Stabilitätskoeffizienten niederschlägt, der mit $v = .989$ nahe am Maximalwert von 1 liegt.

[Tabelle 8 hier einfügen]

Wie aus Tabelle 8 ersichtlich ist, sind die deskriptiven Kennwerte der Indizes aus den Jahren 2012 und 2013 sehr ähnlich. Im Durchschnitt beträgt die mittlere Differenz der Schulindizes $M = .06$ Punkte. Die Ergebnisse des t -Tests bestätigen, dass sich der durchschnittliche Differenzbetrag nicht signifikant von Null unterscheidet ($M = .06$, $SE = .15$, $p > .05$). Diese Ergebnisse gelten sowohl für die Gesamtpopulation (1396 Schulen der Sek. I) als auch für die jeweilige Substichprobe (APS, AHS bzw. Schulen in jedem Bundesland). Die Standardabweichung (SD) der Differenzen beträgt 5.5 Punkte und zeigt, dass für mehr als zwei Drittel der Schulen der ISB zwischen zwei Jahren um weniger als 5.5 Punkte fluktuiert; dies ist geringer als die Hälfte der SD des ISB ($SD = 13.5$, siehe Tabelle 4).

Die Korrelation der paarweisen Schulindizes über zwei Jahre zeigt darüber hinaus eine hohe relative Stabilität. Österreichweit beträgt die Korrelation $r = .92$ ($p < .001$), zwischen den APS ist $r = .91$ ($p < .001$), zwischen den AHS ist $r = .94$ ($p < .001$). In den einzelnen Bundesländern ist dieser Trend ebenfalls zu beobachten, nur im Burgenland fällt die Korrelation mit $r = .68$ ($p < .001$) geringer aus.

Werden die Spalten Min und Max betrachtet, wird allerdings deutlich, dass der Differenzbetrag in manchen Schulen sehr hoch ist. In Oberösterreich gibt es vereinzelt APS mit absoluter Schwankung des ISB größer als 25 Punkte, ca. zwei SD des ISB. Für diese Schulen sind also Daten einer Schülerkohorte nicht auf andere Schülerkohorten übertragbar.

3.5 Der Index und die Schülerleistung

Die Ergebnisse des *multilevel-latent-covariate*-Modells anhand der Daten der VS sind in Tabelle 9 dargestellt.

[Tabelle 9 hier einfügen]

Die VS unterscheiden sich hinsichtlich der aggregierten Schülerleistung ($ICC = .14$). Sowohl auf within-Ebene als auch auf between-Ebene erweist sich der ISB in der VS als signifikanter Prädiktor der Schülerleistung ($\beta_b = -.020$, $p < .001$; $\beta_w = -.010$, $p < .001$). Der Kompositionseffekt wird ebenfalls bestätigt ($\beta_c = -.010$, $p < .001$). Die β -Koeffizienten stellen die durch die Veränderung des ISB um einen Punkt vorhergesagte Leistungsänderung im Standardabweichungsmaß dar. Mit einem Anstieg von 10 ISB-Punkten auf der Schulebene erwartet man modellgemäß eine Verringerung von einem Zehntel der Standardabweichung der Schülerleistung. Durch den ISB können 34% der Leistungsvarianz zwischen den Schulen aufgeklärt werden.

Auf der Sek. I unterscheidet sich die Schülerleistung stärker zwischen Schulen (BIST-Ü M8: $ICC = .33$, BIST-Ü E8: $ICC = .30$). Auch hier gelten die individuellen Merkmale und der ISB der Schulen als bedeutsame Prädiktoren der Schülerleistung sowohl in Englisch ($\beta_b = -.023$, $p < .001$; $\beta_w = -.004$, $p < .001$) als auch in Mathematik ($\beta_b = -.030$, $p < .001$; $\beta_w = -.006$, $p < .001$, siehe Tabelle 10). Auf der Sek. I wird der

Kompositionseffekt des ISB (Modell 1, Tabelle 10) ebenfalls bestätigt, und zwar stärker als in der VS. In Englisch beträgt $\beta_c = -.018$ ($p < .001$), in Mathematik ist $\beta_c = -.024$ ($p < .001$). Zwischen Schulen kann der ISB bis zu 31% der Varianz in Englisch und 50% der Varianz in Mathematik erklären. Insgesamt können die Forschungsfragen 4a und 4b erwartungskonform bestätigt werden.

Nach Kontrolle des Effekts der Schulform (Modell 2, Tabelle 10) verringern sich der Effekt und der Kompositionseffekt des ISB auf Schulebene geringfügig und bleiben weiterhin signifikant. In Englisch wird $\beta_b = -.015$ ($p < .001$), $\beta_c = -.011$ ($p < .001$), in Mathematik wird $\beta_b = -.024$ ($p < .001$), $\beta_c = -.018$ ($p < .001$). Der negative Einfluss des ISB auf die Schulleistung kann demnach größtenteils nicht durch schulformbezogene institutionelle Unterschiede erklärt werden.

[Tabelle 10 hier einfügen]

Insgesamt zeigt sich erwartungsgemäß ein stark negativer Zusammenhang zwischen dem ISB und der Schulleistung, der über den Effekt hinausgeht, der durch den individuellen Hintergrund der einzelnen Schüler/innen bedingt ist. Dies gilt sowohl in der VS als auch auf der Sek. I. Dieser Effekt konfundiert zum großen Teil, anders als in Deutschland, nicht mit den schulformbezogenen Institutionseffekten.

4 Diskussion

In diesem Artikel wurden die Gütekriterien des ISB analysiert und präsentiert. Bedeutsame Unterschiede zwischen Schulen und zwischen Bundesländern hinsichtlich des ISB wurden wie erwartet bestätigt. Die Heterogenität zwischen Schulen ist je nach Bundesland sehr unterschiedlich (siehe Tabellen 3 und 4). In Bundesländern wie dem Burgenland und Kärnten sind Schulen bezüglich ihrer Schülerschaft ziemlich homogen, in Wien hingegen unterscheiden sich Schulen hinsichtlich ihrer Schülerschaft sehr stark. Die Unterschiede zwischen Schulen sind auf der Sek. I größer als auf der Primarstufe. Innerhalb der Sek.-I-Schulen unterscheiden sich die APS stärker voneinander als die AHS.

Insgesamt kann der ISB als ein reliables (robustes) und stabiles Schulmaß angesehen werden, insbesondere auf der Sek. I. Diese Einschätzung gilt jedoch nicht für jede einzelne Schule. In großen Schulen wird der Mittelwert ein guter Repräsentant für die Schülereigenschaften dieser Schulen sein, in der VS ist aufgrund der kleinen Schülerzahl der vierten Schulstufe vieler Schulen der Aggregationswert dagegen erwartungsgemäß weniger reliabel (siehe Tabelle 6). An dieser Stelle sei auf die Arbeit von Kuschej und Schönpflug (2014) hinsichtlich der Berechnung eines Index auf Basis verfügbarer regionaler Daten verwiesen. Praktisch hat jedoch die eingeschränkte Reliabilität des Index in kleinen Schulen wenig Relevanz für die Zuteilung zusätzlicher Mittel, wenn die Betreuungsrelationen (Schülerzahl pro Lehrkraft) wie in dem von Bacher et al. (2010) vorgeschlagenen Mittelzuteilungsmodell mitberücksichtigt werden.

Die Ergebnisse der Stabilitätsanalysen anhand von Daten der Sek.-I-Schulen (Tabelle 8) haben gezeigt, dass der Schulindex mit $v = .989$ über Kohorten hinweg stabil ist. Bei einem Großteil der Schulen ist er ein stabiles Maß (praktisch keine mittlere Veränderung), wenn die absolute Differenz betrachtet wird. Die Korrelation des Schulindex über zwei Messzeitpunkte hinweg (2012 und 2013) zeigt ebenfalls eine hohe Stabilität. Es gibt jedoch Ausnahmen, wenn die Daten einzelner Schulen betrachtet werden. An manchen Schulen kann die

Schülerschaft von Kohorte zu Kohorte einer Schule extrem unterschiedlich sein. Die Reliabilität des ISB könnte erhöht werden, indem Daten mehrerer Schülerkohorten verknüpft werden.

Die Daten der BIST-Ü in Österreich haben den erwarteten Effekt der durch den ISB erfassten Zusammensetzung der Schülerschaft der Schulen auf die Schulleistung in allen getesteten Fächern bestätigt (Tabellen 9 und 10). Auf der Sek. I ist der Effekt etwas stärker ausgeprägt als in der VS. In der BIST-Ü erbringen Schüler/innen in Schulen mit einem höheren ISB in der Regel geringere Leistungen. Darüber hinaus wurde der angenommene Kompositionseffekt der Schulen insofern bestätigt, als die Zusammensetzung der Schülerschaft zusätzlich zu den individuellen Schülermerkmalen einen Einfluss auf die Schülerleistung ausübt. Diese Ergebnisse sprechen dafür, dass der ISB einen Bezug auf unterschiedliche Lehr-Lern-Erfolgchancen der Schulen nimmt, und können als ein Argument für die Validität des ISB als ein Maß des Zusatzbedarfs der Schulen gelten. Anlässlich des Kompositionseffekts des ISB stellt sich allerdings die Frage, ob die vorgeschlagene Berechnungsweise des ISB mittels Aggregation individueller Schülermerkmale zur Bestimmung des Zusatzbedarfs von Schulen optimal ist, weil der Effekt des ISB auf Schülerleistungen stärker ausfällt als der aggregierte Effekt individueller Schülermerkmale.

Die Ergebnisse zeigen weiterhin, dass, anders als in Deutschland, eine starke Konfundierung zwischen dem ISB und der Schulform auf der Sek. I nicht bestätigt wurde: Der Effekt und der Kompositionseffekt des ISB sind auch nach Kontrolle der Schulform signifikant und verringern sich nur geringfügig. Dies deutet auf die Validität des Index als bewährte Einflussgröße der Schulleistung hin.

Zusammengefasst erweist sich der auf BIST-Ü-Daten basierende ISB als ein reliables und stabiles Maß der Schulen mit einem bedeutsamen Bezug zur Schülerleistung. Eine Limitation unseres Beitrags ist allerdings, dass eine Überprüfung des um den Schul- und Unterrichtseffekt bereinigten Zusammenhangs zwischen dem ISB und der Schulleistung nicht möglich ist. Dafür wäre eine Überprüfung des Effekts des ISB unter Kontrolle der anfänglichen Schülerleistung wünschenswert; diese Daten sind leider nicht vorhanden.

Sollte der ISB für eine indexbasierte Verteilung von Ressourcen eingesetzt werden, bedürfen folgende Aspekte noch zusätzlicher Aufmerksamkeit:

a) Die von der BIST-Ü ausgeschlossenen Schüler/innen sind im Index nicht vertreten; ihr Anteil streut stark zwischen Schulen. Diese Schülergruppe wird auch bei der derzeitigen Praxis der Mittelzuweisung an Schulen nur zum Teil berücksichtigt, infolgedessen wird der Zusatzbedarf der Schulen partiell nicht abgedeckt (vgl. Bacher, 2015). Langfristig wäre es wünschenswert, auch die relevanten Variablen ausgeschlossener Schüler/innen erfassen zu können. Kurzfristig müsste der Index im Hinblick auf ausgeschlossene Schüler/innen adjustiert werden. Dazu könnten Daten aus der Bildungsevidenz zur Anzahl der Schüler/innen mit sonderpädagogischem Förderbedarf und außerordentlichem Status verwendet werden.

b) Der ISB repräsentiert inhaltlich den Aspekt des Migrationshintergrunds der Kinder stärker als den Aspekt der sozialen Herkunft. Hier stellt sich die Frage, ob eine alternative Berechnungsweise des Index mit stärkerem Fokus auf die soziale Komposition den Aspekt der Benachteiligung der Schulen besser repräsentieren würde. Dafür können die aktuellen Befunde von Biedermann et al. (2016) über die einzelnen Kompositionseffekte der Schulen und Klassen herangezogen werden.

c) Durch Einbezug anderer Datenquellen, z. B. der Schulstatistik in Form der Bildungsevidenz oder von Daten zum Sozialraum des Arbeitsmarktservice (AMS), sollte der Index besser gegen eine theoretisch mögliche Verfälschbarkeit der Daten (vgl. Bacher, 2015) abgesichert werden.

d) Eine solche Erweiterung der Datenquellen könnte zudem helfen, die Probleme abzumildern, die aufgrund fehlender Daten zur Herkunft der Schüler/innen entstehen, insbesondere, da die Bildungsevidenz jeweils alle vier Schulstufen erfasst und damit auf einer breiteren Grundlage steht.

e) Um den Schulen eine mittelfristige Planung der Ressourcen zu ermöglichen, sollte eine Aktualisierung des Index in längeren Zeiträumen erfolgen. Die Verwendung eines gleitenden Durchschnitts des ISB aus mehreren Erhebungsjahren würde Schwankungen abfedern und zudem die Reliabilität des Index verbessern.

Literatur

Bacher, J. (2015). Indexbasierte Finanzierung des österreichischen Schulsystems. *SchulVerwaltung*, 3(4), 1–4.

Bacher, J. (2016). Mehr Bildungsgerechtigkeit durch sozialindexbasierte Finanzierung des österreichischen Bildungssystems? In H. Seckauer, C. Stelzer-Orthofer & B. Keplinger (Hrsg.): *Das Vorgefunden und das Mögliche* (S. 142–156). Wien: Mandelbaum.

Bacher, J., Altrichter, H., & Nagy, G. (2010). Ausgleich unterschiedlicher Rahmenbedingungen schulischer Arbeit durch eine indexbasierte Mittelverteilung. *Erziehung und Unterricht*, 160(3–4), 384–400.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). *lme4: Linear mixed-effects models using Eigen and S4* (R package version 1.1–8.). <http://CRAN.R-project.org/package=lme4>

Baumert, J., Stanat, P. & Watermann, R. (2006). Schulstruktur und die Entstehung differenzieller Lern- und Entwicklungsmilieus. In J. Baumert, P. Stanat & R. Watermann (Hrsg.), *Herkunftsbedingte Disparitäten im Bildungswesen: Differenzielle Bildungsprozesse und Probleme der Verteilungsgerechtigkeit* (S. 95–188). Wiesbaden: VS Verlag für Sozialwissenschaften.

Biedermann, H., Weber, C., Herzog-Punzenberger, B. & Nagel, A. (2016). Auf die Mitschüler/innen kommt es an? Schulische Segregation – Effekte der Schul- und Klassenzusammensetzung auf der Primarstufe und der Sekundarebene I. In M. Bruneforth, F. Eder, K. Krainer, C. Schreiner, A. Seel & C. Spiel (Hrsg.), *Nationaler Bildungsbericht 2015 – Band 2. Fokussierte Analysen bildungspolitischer Schwerpunktthemen* (S. 133–173). Graz: Leykam.

BIFIE. (2015). *BIFIEsurvey: Some tools for survey statistics in educational assessment, developed by BIFIE*. <http://CRAN.R-project.org/package=BIFIEsurvey>

- Breit, S., & Schreiner, C. (2006). Sozioökonomische Herkunft und Schulleistung. In G. Haider & C. Schreiner (Hrsg.), *Die PISA-Studie: Österreichs Schulsystem im internationalen Wettbewerb* (S. 195–210). Wien: Böhlau.
- Bruneforth, M., & Lassnigg, L. (2012). *Nationaler Bildungsbericht Österreich 2012, Band 1: Das Schulsystem im Spiegel von Daten und Indikatoren*. Graz: Leykam.
- Bruneforth, M., Weber, C., & Bacher, J. (2012). Chancengleichheit und garantiertes Bildungsminimum in Österreich. In B. Herzog-Punzenberger (Hrsg.), *Nationaler Bildungsbericht Österreich 2012. Band 2. Fokussierte Analysen bildungspolitischer Schwerpunktthemen* (S. 187–226). Graz: Leykam.
- Dumont, H., Neumann, M., Maaz, K., & Trautwein, U. (2013). Die Zusammensetzung der Schülerschaft als Einflussfaktor für Schulleistungen: Internationale und nationale Befunde. *Psychologie in Erziehung und Unterricht*, 60, 163–183.
- Fazekas, M. (2012). School funding formulas: Review of main characteristics and impacts. *OECD Education Working Papers*, No. 74. Paris: OECD Publishing. <http://dx.doi.org/10.1787/5k993xw27cd3-en>
- Ganzeboom, H. B. G., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56.
- George, A. C., & Robitzsch, A. (2015). „Fairer Vergleich“: *Technische Dokumentation – BIST-Ü Mathematik, 4. Schulstufe, 2013*. Salzburg: BIFIE. www.bifie.at/node/2766
- Gröhlich, C., Guill, K., Scharenberg, K. & Bos, W. (2010). Differenzielle Lern- und Entwicklungsmilieus beim Erwerb der Lesekompetenz in den Jahrgangsstufen 7 und 8. In W. Bos & C. Gröhlich (Hrsg.), *KESS 8 – Kompetenzen und Einstellungen von Schülerinnen und Schülern am Ende der Jahrgangsstufe 8* (S. 100–106). Münster: Waxmann.
- Harrell, F. E. (2015). *Hmisc: Harrell Miscellaneous*. <http://CRAN.R-project.org/package=Hmisc>
- Kuscheij, H., & Schönpflug, K. (2014). *Indikatoren bedarfsorientierter Mittelverteilung im österreichischen Pflichtschulwesen*. Wien: AK Wien.
- Levačić, R. (2008). Funding schools by formula. In N. C. Soguel & P. Jaccard (Hrsg.), *Governance and performance of education systems* (S. 205–235). Dordrecht: Springer.

- Lüdtke, O., Marsh, H. W., Robitzsch, A., Trautwein, U., Asparouhov, T., & Muthén, B. (2008). The multilevel latent covariate model: A new, more reliable approach to group-level effects in contextual studies. *Psychological Methods*, 13(3), 203–229.
- Lüdtke, O., & Robitzsch, A. (2010). Umgang mit fehlenden Daten in der empirischen Bildungsforschung. In S. Maschke & L. Stecher (Hrsg.), *Enzyklopädie Erziehungswissenschaft Online. Fachgebiet Methoden der empirischen erziehungswissenschaftlichen Forschung, Quantitative Forschungsmethoden* (S. 1–42). Weinheim: Juventa.
- Maaz, K., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes. *Child Development Perspectives*, 2(2), 99–106.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149–174.
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus user's guide* (7th ed.). Los Angeles: Muthén & Muthén.
- Pham, G., & Robitzsch, A. (2015). „Fairer Vergleich“: Technische Dokumentation – BIST-Ü Englisch, 8. Schulstufe, 2013. Salzburg: BIFIE. www.bifie.at/node/2848
- Pornprasertmanit, S., Miller, P., Schoemann, A., & Rosseel, Y. (2013). *semTools: Useful tools for structural equation modeling*. <http://CRAN.R-project.org/package=semTools>
- R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. www.R-project.org
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Nielsen & Lydiche.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models*. Thousand Oakes: Sage.
- Revelle, W. (2014). *psych: Procedures for psychological, psychometric, and personality research*. Evanston, Illinois: Northwestern University. <http://CRAN.R-project.org/package=psych>
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63, 581–592.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.

- Schreiner, C. & Breit, S. (Hrsg.). (2012). *Standardüberprüfung 2012 Mathematik, 8. Schulstufe – Bundesergebnisbericht*. Salzburg: BIFIE. www.bifie.at/node/1948
- Schreiner, C. & Breit, S. (Hrsg.). (2014a). *Standardüberprüfung 2013 Mathematik, 4. Schulstufe – Bundesergebnisbericht*. Salzburg: BIFIE. www.bifie.at/node/2489
- Schreiner, C. & Breit, S. (Hrsg.). (2014b). *Standardüberprüfung 2013 Englisch, 8. Schulstufe – Bundesergebnisbericht*. Salzburg: BIFIE. www.bifie.at/node/2490
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London: Sage Publishers.
- Teddlie, C., Reynolds, D., & Sammons, P. (2000). The methodology and scientific properties of school effectiveness research. In C. Teddlie & D. Reynolds (Hrsg.), *The international handbook of school effectiveness research* (pp. 55–133). New York: Routledge.
- Zhao, J. H., & Schafer, J. L. (2013). *pan: Multiple imputation for multivariate panel or clustered data*. <https://cran.r-project.org/package=pan>

Tabellen

Tabelle 1: Teilnehmerstatistik in den Bildungsstandardüberprüfungen (BIST-Ü)

<i>Schülerstatistik</i>		<i>M4 2013</i>	<i>M8 2012</i>	<i>E8 2013</i>
<i>Gemeldete Gesamtschülerzahl (100%)</i>		79266	86868	83745
Schüler/innen mit Körper- oder Sinnesbehinderung	Anzahl	191	152	164
	Quote	0.2%	0.2%	0.2%
Außerordentliche Schüler/innen	Anzahl	586	756	843
	Quote	0.7%	0.9%	1.0%
Lehrplan der allgemeinen Sonderschule (fachspezifisch) oder einer niedrigeren Schulstufe	Anzahl	2149	1789	1743
	Quote	2.7%	2.1%	2.1%
Gesamtzahl ausgeschlossener Schülerinnen und Schüler	Anzahl	2926	2697	3541
	Quote	3.7%	3.1%	3.3%
Schulwechsel, Schulabbruch	Anzahl	543	790	740
	Quote	0.7%	0.9%	0.9%
<i>Zielpopulation der BIST-Ü (100%)</i>		75797	83381	80204
Teilnehmerzahl	Anzahl	73655	79678	76728
	Quote	97.2%	95.6%	95.7%

Quelle: Schreiner und Breit, 2012, 2014a, 2014b

Anmerkung: M4 2013 = Fach Mathematik, Schulstufe 4, Erhebungsjahr 2013; M8 2012 = Fach Mathematik, Schulstufe 8, Erhebungsjahr 2012; E8 2013 = Fach Englisch, Schulstufe 8, Erhebungsjahr 2013.

Tabelle 2: Missinganteile (%) der Daten auf der Ebene der Schüler/innen

Variable	Erhebungsfach, Schulstufe, Jahr		
	Mathematik, 4, 2013	Mathematik, 8, 2012	Englisch, 8, 2013
Migrationshintergrund	1.4	1.0	0.9
Muttersprache	1.8	4.8	2.2
Elternausbildung	7.6	11.0	11.5
HISEI	4.9	4.9	4.8

Tabelle 3: Statistik des Schulindex der Volksschulen (2013)

Stichprobe	N	<i>M</i>	<i>SD</i>	<i>Q5</i>	<i>Q25</i>	<i>Q75</i>	<i>Q95</i>	ICC
Burgenland	178	109.7	8.4	100.0	103.7	113.3	125.0	.05
Kärnten	239	108.5	5.9	100.0	104.5	111.1	119.4	.07
Niederösterreich	609	111.0	8.3	101.5	105.7	114.1	126.6	.10
Oberösterreich	559	112.0	10.8	101.5	105.3	114.6	136.9	.17
Salzburg	184	111.4	9.4	100.0	105.1	114.4	130.1	.14
Steiermark	479	110.2	9.8	100.0	104.5	112.5	127.4	.19
Tirol	377	111.3	8.8	100.0	105.0	115.8	127.8	.08
Vorarlberg	163	112.6	9.7	100.0	104.4	118.0	130.7	.08
Wien	260	128.2	16.3	105.7	113.5	142.7	154.2	.28
Gesamt	3048	112.4	11.1	100.0	105.3	115.5	137.3	.18

Anmerkung: N = Anzahl Schulen, *M* = Mittelwert, *SD* = Standardabweichung, *Q5* = 5%-Quantil, *Q25* = 25%-Quantil, *Q75* = 75%-Quantil, *Q95* = 95%-Quantil, ICC = Intraklassenkorrelation (Intraclass Correlation Coefficient).

Tabelle 4: Statistik des Schulindex der Sekundarstufe I-Schulen (2013)

Stichprobe	N	M	SD	Q5	Q25	Q75	Q95	ICC
APS	1143	118.9	13.8	106.0	109.9	122.7	151.9	.24
AHS	267	110.1	9.3	102.1	104.6	112.0	129.8	.19
Burgenland	49	113.2	6.3	104.1	108.4	117.9	124.5	.06
Kärnten	85	113.4	7.9	104.0	107.5	116.3	128.3	.11
Niederösterreich	300	113.4	8.3	103.4	108.2	116.5	128.2	.12
Oberösterreich	277	116.2	13.1	103.9	107.8	118.9	147.1	.25
Salzburg	92	115.3	12.0	102.5	107.8	117.0	136.3	.21
Steiermark	208	114.9	11.7	104.1	108.2	116.2	138.8	.24
Tirol	129	114.7	8.7	104.9	108.8	117.6	131.8	.11
Vorarlberg	67	116.9	10.1	103.1	108.8	123.5	135.8	.12
Wien	203	131.8	19.0	105.3	113.3	148.5	161.1	.34
Gesamt	1410	117.2	13.5	104.1	108.4	120.7	149.6	.26

Anmerkung: N = Anzahl Schulen, M = Mittelwert, SD = Standardabweichung, Q5 = 5%-Quantil, Q25 = 25%-Quantil, Q75 = 75%-Quantil, Q95 = 95%-Quantil, ICC = Intraklassenkorrelation (Intraclass Correlation Coefficient).

Tabelle 5: Deskriptive Statistik und Korrelationen der vier Indikatoren des ISB (2013)

Stufe	Variable	<i>M</i>	<i>SD</i>	Migrations- hintergrund	Mutter- sprache	Eltern- ausbildung	HISEI	$\bar{B}_j$	B_{ij}
4	Migrationshintergrund	.12	.18		.69	.28	.25	.91	.82
	Muttersprache	.11	.16	.93		.25	.21	.90	.79
	Elternausbildung	.06	.08	.52	.51		.22	.68	.54
	HISEI	.21	.15	.22	.21	.31		.56	.63
8	Migrationshintergrund	.18	.21		.77	.21	.24	.94	.81
	Muttersprache	.17	.20	.97		.22	.22	.95	.80
	Elternausbildung	.13	.10	.61	.63		.20	.79	.54
	HISEI	.22	.12	.38	.40	.62		.63	.62

Anmerkung: *M* = Mittelwert, *SD* = Standardabweichung des Schulindex. Die Spalten 5–8 enthalten die Korrelationen der vier Indikatoren des Index miteinander, oberhalb der Diagonale sind Korrelationen anhand der Daten auf Individualebene, unterhalb der Diagonale sind Korrelationen anhand der Daten auf Schulebene. Die Korrelationen zwischen den Indikatoren und dem Schulindex sind in Spalte $\bar{B}_j$ und zwischen den individuellen Indikatoren und den Schülermerkmalen in B_{ij} dargestellt.

Tabelle 6: Schulgrößen und Reliabilität des ISB der Volksschulen (2013)

Stichprobe	$\bar{n}$	ICC(2, $\bar{n}$)	ICC(2,20)	ICC(2,50)
Burgenland	13	.41	.52	.73
Kärnten	19	.59	.60	.79
Niederösterreich	24	.73	.69	.85
Oberösterreich	24	.83	.81	.91
Salzburg	27	.81	.76	.89
Steiermark	21	.83	.82	.92
Tirol	17	.60	.63	.81
Vorarlberg	23	.65	.62	.80
Wien	53	.95	.89	.95
Gesamt	24	.84	.82	.92

Anmerkung: $\bar{n}$ = durchschnittliche Schulgröße, ICC(2, $\bar{n}$) = ICC(2) bei durchschnittlicher Schulgröße, ICC(2,20) = ICC(2) bei einer Schulgröße von 20, ICC(2,50) = ICC(2) bei einer Schulgröße von 50.

Tabelle 7: Schulgrößen und Reliabilität des ISB auf der Sekundarstufe I (2013)

Stichprobe	$\bar{n}$	ICC(2, $\bar{n}$)	ICC(2,20)	ICC(2,50)
APS	45	.93	.86	.94
AHS	94	.96	.82	.92
Burgenland	50	.77	.57	.77
Kärnten	60	.88	.71	.86
Niederösterreich	48	.87	.74	.88
Oberösterreich	50	.94	.87	.94
Salzburg	58	.94	.84	.93
Steiermark	51	.94	.86	.94
Tirol	53	.87	.71	.86
Vorarlberg	58	.89	.74	.88
Wien	69	.97	.91	.96
Gesamt	54	.95	.87	.95

Anmerkung: $\bar{n}$ = durchschnittliche Schulgröße, ICC(2, $\bar{n}$) = ICC(2) bei durchschnittlicher Schulgröße, ICC(2,20) = ICC(2) bei einer Schulgröße von 20, ICC(2,50) = ICC(2) bei einer Schulgröße von 50.

Tabelle 8: Deskriptiva des Schulindex der Sekundarstufe I-Schulen (Erhebungsjahre 2012, 2013)

Schulstichprobe	N	ISB (2012)			ISB (2013)			Differenz (2013–2012)					
		<i>M</i>	<i>SD</i>	$\bar{n}$	<i>M</i>	<i>SD</i>	$\bar{n}$	<i>M</i>	<i>SE</i>	<i>SD</i>	<i>r</i>	Min	Max
Gesamt	1396	117.2	13.5	57	117.3	13.6	55	0.06	0.15	5.5	.92	-25.8	48.0
APS	1130	119.0	13.7	47	119.0	13.9	46	-0.01	0.18	5.9	.91	-25.8	48.0
AHS	266	109.8	9.3	97	110.2	9.3	94	0.35	0.20	3.3	.94	-16.3	12.3
Burgenland	49	112.8	6.9	53	113.2	6.3	50	0.47	0.78	5.3	.68	-14.1	11.8
Kärnten	82	113.5	8.2	64	113.5	7.9	61	0.05	0.56	5.0	.81	-16.8	13.4
Niederösterreich	297	113.9	8.5	50	113.4	8.2	48	-0.58	0.33	5.6	.78	-18.1	19.5
Oberösterreich	274	116.1	12.7	53	116.3	13.1	51	0.22	0.38	6.3	.88	-25.8	48.0
Salzburg	91	115.7	11.7	60	115.4	12.0	59	-0.37	0.55	5.1	.91	-12.5	20.6
Steiermark	205	114.4	11.6	53	114.9	11.7	51	0.48	0.32	4.5	.93	-14.9	13.3
Tirol	128	115.3	9.4	55	114.8	8.7	54	-0.50	0.41	4.6	.88	-15.6	10.7
Vorarlberg	67	116.0	9.1	61	116.9	10.1	58	0.85	0.74	5.9	.81	-13.8	22.5
Wien	203	131.3	19.4	70	131.8	19.0	69	0.54	0.40	5.7	.96	-16.9	15.1

Anmerkung: N =Anzahl Schulen, *M* = Mittelwert, *SD* = Standardabweichung des Schulindex, $\bar{n}$ = durchschnittliche Schulgröße.

Tabelle 9: Ergebnisse des multilevel-latent-covariate-Modells (Schulstufe 4)

Fach Mathematik, 2013		
Feste Effekte	Koeffizient	SE
Intercept	-.013*	.007
ISB between	-.020***	.001
ISB within	-.010***	.000
ISB Kompositionseffekt	-.010***	.001
Zufällige Effekte	Varianz	SE
Varianz between	.087	.004
Varianz within	.807	.005
Aufgeklärte Varianz	R ²	SE
R ² between	.340	.020
R ² within	.055	.002

Anmerkung: Signifikanzniveau: $p < .05$ **, $p < .001$ ***.

Tabelle 10: Ergebnisse des multilevel-latent-covariate-Modells (Schulstufe 8)

	Fach Mathematik, 2012				Fach Englisch, 2013			
	Modell 1		Modell 2		Modell 1		Modell 2	
	Koeffizient	SE	Koeffizient	SE	Koeffizient	SE	Koeffizient	SE
Feste Effekte								
Intercept	-.078***	.011	-.229***	.010	-.110***	.012	-.303***	.009
Schulform	-	-	.708***	.021	-	-	.937***	.017
ISB between	-.030***	.001	-.024***	.001	-.023***	.001	-.015***	.001
ISB within	-.006***	.000	-.006***	.000	-.004***	.000	-.004***	.000
ISB Kompositionseffekt	-.024***	.001	-.018***	.001	-.018***	.001	-.011***	.001
Zufällige Effekte	Varianz	SE	Varianz	SE	Varianz	SE	Varianz	SE
Varianz between	.156	.007	.078	.004	.193	.008	.058	.003
Varianz within	.620	.005	.620	.005	.647	.005	.648	.005
Aufgeklärte Varianz	R^2	SE	R^2	SE	R^2	SE	R^2	SE
R^2 between	.502	.020	.746	.013	.311	.021	.784	.013
R^2 within	.031	.000	.031	.001	.013	.001	.013	.001

Anmerkung: Modell 1: nur ISB als Prädiktor, Modell 2: ISB als Prädiktor und Schulform zusätzlich als Prädiktor auf Schulebene. Signifikanzniveau: $p < .001$ '***'.