

Schul- und Klassenzusammensetzung: Methodisches Vorgehen und Zusatzergebnisse

Onlinedokument zum Beitrag „Wer unterrichtet wen? Die Zusammensetzung von Klassen und Schulen und die Zuteilung von Lehrkräften“ im Band 2 des Nationalen Bildungsberichts Österreich 2018).

<http://doi.org/10.17888/nbb2018-2-4-2>

Christoph Weber

Einleitung

In diesem Dokument werden die methodische Vorgehensweise sowie die Detailergebnisse in Bezug auf das Kapitel „Wer unterrichtet wen? Die Zusammensetzung von Klassen und Schulen und die Zuteilung von Lehrkräften“ im Band 2 des Nationalen Bildungsberichts (NBB) 2018 (Weber, Moosbrugger, Hasengruber, Altrichter & Schrodtr, 2019) dargestellt. Vorab soll darauf verwiesen werden, dass je nach Fragestellung unterschiedliche Daten herangezogen wurden. Es wurde auf Daten der Bildungsstandardüberprüfung auf der 4. Schulstufe in Mathematik (BIST-Ü-M4 2013; vgl. BIFIE, 2015) und Deutsch (BIST-Ü-D4 2015; BIFIE, 2017a) und auf der 8. Schulstufe in Deutsch (BIST-Ü-D8 2016; vgl. BIFIE, 2017b) sowie TALIS 2008 (OECD, 2010) zurückgegriffen.

A Effekte der Schul- und Klassenkomposition auf die Leseleistung

In Biedermann, Weber, Herzog-Punzenberger und Nagel (2016a) wurde der Einfluss der Schul- und Klassenkomposition auf die Mathematikleistungen untersucht. Für den aktuellen Beitrag wurden diese Analysen mit den Daten der BIST-Ü-D4 2015 und BIST-Ü-D8 2016 wiederholt, wobei an dieser Stelle nur auf Lesen fokussiert wird. Für die Sekundarstufe wurden getrennte Analysen für AHS (ca. 34% der getesteten Schüler/innen) und NMS (ca. 42 % der getesteten Schüler/innen) durchgeführt. Analysen für die auslaufende Hauptschule wurden nicht durchgeführt (24%), weil hier u.a. die Analyse von Lerngruppeneffekten durch die Leistungsgruppierung erschwert wird. Nachfolgend werden die Analysemodelle, die Behandlung der fehlenden Werte und die Ergebnisse kurz dargestellt.

Analysemodelle

Die Analyse der Kompositionseffekte basiert auf einem Dreiebenenregressionsmodell (Ebene 1/L1 = Schüler/innen; Ebene 2/L2 = Klassen und Ebene 3/L3 = Schulen) der Form:

$$\text{Lesen}_{ijk} = \beta_0 + \beta_1(X_{ijk}) + \beta_2(X_{.jk}) + \beta_3(X_{..k}) + e_{ijk} + u_{0jk} + v_{00k}$$

X_{ijk} ist ein L1-Merkmal (z.B. Sozialstatus einer Schülerin/eines Schülers), $X_{.jk}$ das auf L2 aggregierte L1-Merkmal (z.B. der durchschnittliche Sozialstatus einer Klasse), und $X_{..k}$ ist das auf L3 aggregierte L1-Merkmal (z.B. der durchschnittliche Sozialstatus einer Schule). Für die Analysen wurden die Merkmale Sozialstatus, nichtdeutsche Alltagssprache und Geschlecht (inkl. ihrer Kompositionseffekte auf L2 und L3 berücksichtigt). Auf L1 wurde zusätzlich das Alter der Schüler/innen (in z-standardisierter Form) kontrolliert. Die L1- und L2-Variablen wurden am Gesamtmittelwert (grandmean) zentriert (vgl. Enders & Tofighi, 2007). L2- und L3-Effekte bilden somit direkt Kompositionseffekte ab. Die Analysen wurden mit Mplus 7.3 (Muthén & Muthén, 1998-2017) unter Verwendung einer robusten Maximum Likelihood Schätzung (MLR) durchgeführt. Eine nähere Diskussion der Analysemodelle findet sich in Biedermann, Weber, Herzog-Punzenberger und Nagel (2016b).

Zur Beurteilung der Stärke der Kompositionseffekte wurden Effektstärken nach Marsh et al. (2012) als $ES=2b(SD_{\text{pred}})/(SD_{\text{uwl1}})$ berechnet, wobei b der unstandardisierte Kontexteffekt, SD_{pred} die Standardabweichung des Prädiktors auf L2 bzw. L3 und SD_{uwl1} die L1-Standardabweichung der abhängigen Variable Lesen ist. ES ist vergleichbar mit Cohens d (vgl. u.a. Cohen, 1992) und wird in Anlehnung an Eder, Altrichter, Hofmann und Weber (2015) wie folgt interpretiert: Effektstärken ab 0,15 werden als schwach, ab 0,35 als mittel und ab 0,55 als stark bezeichnet.

Behandlung fehlender Werte

Für eine adäquate Behandlung der fehlenden Werte bei den Daten der Schüler/innen- und Elternfragebögen wurde eine multiple Imputation mit der Software Blimp 1 (Enders, Keller & Levy, 2017; Keller & Enders, 2017) durchgeführt. Blimp implementiert einen FCS (fully conditional specification) Ansatz, mit dem die

Mehrebenenstruktur (Random Intercepts bis zu drei Ebenen) berücksichtigt werden kann. Auch ist es möglich, kategoriale Variablen (ordinal, nominal) in den Modellen zu verwenden. Da für die abhängige Variable „Lesen“ Plausible Werte (10 PVs) vorliegen, wurden getrennte Imputationen für die einzelnen PVs durchgeführt (nested imputations, siehe z.B. Kaplan & Su, 2018). D.h.: Im ersten Imputationsmodell wurde PV1 verwendet, im zweiten Imputationsmodell PV2, usw.... Pro Imputationsmodell wurden 10 imputierte Datensätze erzeugt. Somit ergeben sich in der Summe 100 imputierte Datensätze. Da genestete Imputationen korreliert sind, müssten die üblichen Regeln zur Berechnung der kombinierten Standardfehler adaptiert werden (Rubin, 2003). In Mplus (Muthén & Muthén, 1998-2017) - das für die anschließenden Analysen verwendet wurde – ist die Korrektur nicht implementiert. Da aufgrund der großen Stichprobe inhaltlich substantielle Effekte (gemessen an der Effektstärke, siehe unten) statistisch hoch signifikant sind ($p < .001$) und nicht davon auszugehen ist, dass sich durch die Korrektur in Hinblick auf die statistische Signifikanz inhaltlich substantieller Effekte etwas ändert, wurde auf eine „händische“ Korrektur verzichtet.

Im Detail wurden getrennte Dreiebenen¹-Imputationsmodelle für VS, AHS und NMS jeweils auch noch differenziert nach dreistufigem Urbanisierungsgrad (städtisch, intermediär und ländlich) gerechnet. Verwendet wurden nur Schulen mit mindestens zwei Klassen und mindestens zehn Schüler/innen je Klasse (Marsh, Kuyper, Morin, Parker & Seaton, 2014). Für die Imputationen wurden alle Variablen des Analysemodells berücksichtigt (siehe unten), wobei die Indikatoren des Sozialstatus (Anzahl Bücher, Bildung der Eltern und HISEI; siehe dazu Freunberger, Robitzsch & Pham, 2014) einzeln in das Imputationsmodell aufgenommen wurden. Darüber hinaus wurden auch noch weitere Variablen im Imputationsmodell berücksichtigt (auf Schulebene: Privatschule; auf Schüler/innenebene: PVs der anderen Kompetenzbereiche in Deutsch, Migrationshintergrund und die Angaben zum Bemühen bei den BIST-Ü im Vergleich zu einer Schularbeit). Durch die Verwendung von Klassen- und Schulmittelwerten wurden bei der Imputation explizit Kontexteffekte berücksichtigt (Keller & Enders, 2017). Bei der Imputation wurden mind. 10.000 burn-in Iterationen und ein Thinning-Interval von 5.000 verwendet. Die Wahl der burn-in Iterationen und des Thinning-Intervals wurden auf Basis des PSR-Wert (potential scale reduction), der zur Diagnose der Konvergenz des MCMC-Algorithmus in Blimp implementiert ist, spezifiziert.

Ergebnisse

Die Ergebnisse spiegeln weitgehend die Befunde von Biedermann et al. (2016a) wider, auch wenn die ES zum Teil etwas schwächer ausfallen. Auf Klassenebene zeigen sich Effekte des durchschnittlichen Sozialstatus und auf Schulebene signifikante Effekte des Anteils der Schüler/innen mit nichtdeutscher Muttersprache.

Tabelle 1: Ergebnisse Dreiebenenmodelle - Kompositionseffekte für Lesen

	VS			NMS			AHS		
	b	SE	ES	b	SE	ES	b	SE	ES
<i>Schülerinnenebene</i>									
SES	48,43***	0,63		24,83***	0,78		21,29***	0,81	
Geschlecht männlich	-26,98***	0,80		-23,96***	1,19		-20,11***	1,23	
Nichtdeutsche Alltagssprache	-28,12***	1,09		-32,52***	1,55		-31,49***	1,17	
Alter	-11,24***	0,43		-24,67***	1,03		-11,35***	1,17	
<i>Klassenebene</i>									
durchschnittlicher SES	19,67***	3,19	0,18	15,82***	3,93	0,13	22,85***	4,44	0,21
%männlich	-8,61	6,78	-0,02	-12,52	6,38	-0,04	-23,83***	4,96	-0,12
%nichtdeutsche Alltagssprache	-7,24	6,87	-0,03	-22,65**	7,57	-0,13	-11,73	10,75	-0,05
<i>Schulebene</i>									
durchschnittlicher SES	-4,75	3,95	-0,04	-11,44	13,20	-0,08	-11,51	6,33	-0,09
%männlich	-3,13	10,80	-0,02	-7,91	6,85	-0,02	8,66	10,59	0,03
%nichtdeutsche Alltagssprache	-19,25*	7,84	-0,08	-43,93***	10,40	-0,24	-28,55*	13,73	-0,12
<i>Varianzen</i>									
Schüler/innen	6570,53	50,50		5651,68	65,66		5180,48	63,61	
Klassen	332,22	26,43		272,59	34,00		280,62	31,11	
Schule	282,97	27,82		503,10	49,42		243,36	30,50	
Intercept	537,48			877,67			780,74		
<i>N</i>									
Schüler/innen	54297			28937			24791		
Klassen	3003			1589			1127		
Schulen	1261			598			264		

¹ Schüler/in, Klasse und Schule

B Kriterien der Klassenzusammenstellung und innerschulische Segregation – Lassen sich Hinweise auf eine segregierende Wirkung einzelner Zusammenstellungskriterien finden?

In Abschnitt 2.3. wurde der Frage nachgegangen, ob Zusammenhänge zwischen unterschiedlichen Kriterien der Klassenzusammenstellung und der innerschulischen Segregation bestehen. Informationen zu Kriterien der Klassenzusammenstellung liegen auf Basis der Schulleiterangaben im Rahmen der BISTÜ-M413 (BIFIE, 2015) und BISTÜ-D816 (BIFIE, 2017b) vor.

Analysemodelle

Nachfolgend wird die Vorgehensweise zur Beantwortung der Frage nach einem Zusammenhang zwischen innerschulischer Segregation und Kriterien der Klassenzusammenstellung beschrieben.

Die Schätzung der innerschulischen Segregation erfolgt auf Basis von Dreiebenenmodellen für den Sozialstatus (soziale Segregation) und den Migrationshintergrund (ethnische Segregation). Für das kontinuierliche Merkmal Sozialstatus wurde ein Modell der Form

$$\begin{aligned} \text{Sozialstatus}_{ijk} &= \beta_0 + e_{ijk} + u_{jk} + v_k \\ e_{ijk} &\sim N(0, \sigma_e^2), u_{jk} \sim N(0, \sigma_u^2), v_k \sim N(0, \sigma_v^2) \end{aligned}$$

geschätzt. σ_e^2 ist die Residualvarianz auf Schülerebene, σ_u^2 ist die Varianz auf Klassenebene, und σ_v^2 ist die Varianz auf Schulebene. Die Summe $\sigma_u^2 + \sigma_e^2$ ist die Varianz des Sozialstatus innerhalb von Schulen. Als Maßzahl der innerschulischen Segregation wurde nun eine Variante der Intraklassenkorrelation (ICC) wie folgt berechnet:

$$ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

ICC gibt somit den Anteil der innerschulischen Varianz des Sozialstatus wieder, der durch die Klassenzugehörigkeit erklärt werden kann. Unter Verwendung dieser Maßzahl für die innerschulische soziale Segregation wurden Gruppenvergleiche durchgeführt², wobei für Schulen, bei denen laut Angabe der Schulleitungen ein Zusammenstellungskriterium entscheidend (Gruppe 1), etwas entscheidend (Gruppe 2) bzw. gar nicht entscheidend (Gruppe 3) war, die Modellparameter getrennt geschätzt wurden:

$$\begin{aligned} \text{Sozialstatus}_{ijk}^{G1} &= \beta_0^{G1} + e_{ijk}^{G1} + u_{jk}^{G1} + v_k^{G1} \\ \text{Sozialstatus}_{ijk}^{G2} &= \beta_0^{G2} + e_{ijk}^{G2} + u_{jk}^{G2} + v_k^{G2} \\ \text{Sozialstatus}_{ijk}^{G3} &= \beta_0^{G3} + e_{ijk}^{G3} + u_{jk}^{G3} + v_k^{G3} \\ e_{ijk}^{G1} &\sim N(0, \sigma_{eG1}^2), u_{jk}^{G1} \sim N(0, \sigma_{uG1}^2), v_k^{G1} \sim N(0, \sigma_{vG1}^2) \\ e_{ijk}^{G2} &\sim N(0, \sigma_{eG2}^2), u_{jk}^{G2} \sim N(0, \sigma_{uG2}^2), v_k^{G2} \sim N(0, \sigma_{vG2}^2) \\ e_{ijk}^{G3} &\sim N(0, \sigma_{eG3}^2), u_{jk}^{G3} \sim N(0, \sigma_{uG3}^2), v_k^{G3} \sim N(0, \sigma_{vG3}^2) \end{aligned}$$

² Da ein Gruppenvergleich unter der Verwendung des *grouping* Befehls bei Dreiebenenmodellen in Mplus 7.3 nicht möglich ist, wurde eine alternative Modellspezifikation verwendet, die identische Ergebnisse liefert (verglichen mit Ergebnissen von getrennten Analysen für Subgruppen). Bei dem Ansatz werden auf Basis der Variable Sozialstatus drei Variablen Sozialstatus^{G1}, Sozialstatus^{G2} und Sozialstatus^{G3} erzeugt: Sozialstatus^{G1} = Sozialstatus wenn G = 1 und Sozialstatus^{G1} = missing, wenn G = 2, oder G = 3. Sozialstatus^{G2} = Sozialstatus wenn G = 2 und Sozialstatus^{G2} = missing, wenn G = 1, oder G = 3, usw.... Anschließend werden alle drei Variablen simultan im Modell berücksichtigt. Für jede Variable wird auf allen Ebenen eine eigene Varianzkomponente und auf L3 ein Intercept geschätzt. Da je Fall nur gültige Werte für eine Variable vorliegen und somit der Kovarianzteil des Modells (Cov(Sozialstatus^{G1}, Sozialstatus^{G2}), ...) empirisch nicht bestimmt ist, werden diese Kovarianzen auf 0 fixiert.

Mittels Wald-Test wurde anschließend geprüft, ob sich das Ausmaß der innerschulischen sozialen Segregation der Schulen der Gruppe 1 von Schulen der Gruppe 3 unterscheiden. Formal wurde folgende Modellrestriktion (H0) getestet:

$$ICC^{G1} = \frac{\sigma_{uG1}^2}{\sigma_{uG1}^2 + \sigma_{eG1}^2} = ICC^{G3} = \frac{\sigma_{uG3}^2}{\sigma_{uG3}^2 + \sigma_{eG3}^2}$$

Ein signifikantes Ergebnis des Wald-Tests weist darauf hin, dass H0 abzulehnen ist, sich also die innerschulische Segregation zwischen Schulen der Gruppe 1 und Gruppe 3 unterscheidet. Die Modellschätzung erfolgt mittels MLR-Schätzung.

Für das Merkmal Migrationshintergrund (nein/ja) wurde ein Dreiebenen-Binomial-Response-Modell (Leckie et al., 2012) mittels Bayes-Schätzung (default priors; Asparouhov & Muthén, 2010) in Mplus geschätzt:

$$\begin{aligned} y_{jk} &\sim \text{Binomial}(n_{jk}, \pi_{jk}) \\ \text{probit}(\pi_{jk}) &= \beta_0 + u_{jk} + v_k \\ u_{jk} &\sim N(0, \sigma_u^2), v_k \sim N(0, \sigma_v^2) \end{aligned}$$

y_{jk} ist dabei der Anteil der Schüler/innen mit Migrationshintergrund in der Klasse j der Schule k. n_{jk} ist die Anzahl der Schüler/innen der Klasse j in Schule k und π_{jk} ist der unbekannte, der Klasse j (aus Schule k) zugrundeliegende Anteil der Schüler/innen mit Migrationshintergrund. π_{jk} wird über eine Linkfunktion – in diesem Fall ein probit-Link³ – mit dem linearen Prädiktor $\beta_0 + u_{jk} + v_k$ in Verbindung gesetzt. Die Varianz σ_v^2 bildet zwischenschulische Unterschiede im Migrationsanteil ab und σ_u^2 stellt innerschulische Unterschiede (d.h. Abweichungen der Klassenanteile vom Schulanteil) dar und wird hier als eine direkte Maßzahl der innerschulischen Segregation verwendet (siehe auch Leckie et al., 2012). Analog zum Sozialstatus wurde ein Gruppenvergleich gerechnet⁴:

$$\begin{aligned} \text{probit}(\pi_{jk}^{G1}) &= \beta_0^{G1} + u_{jk}^{G1} + v_k^{G1} \\ \text{probit}(\pi_{jk}^{G2}) &= \beta_0^{G2} + u_{jk}^{G2} + v_k^{G2} \\ \text{probit}(\pi_{jk}^{G3}) &= \beta_0^{G3} + u_{jk}^{G3} + v_k^{G3} \\ u_{jk}^{G1} &\sim N(0, \sigma_{uG1}^2), v_k^{G1} \sim N(0, \sigma_{vG1}^2) \\ u_{jk}^{G2} &\sim N(0, \sigma_{uG2}^2), v_k^{G2} \sim N(0, \sigma_{vG2}^2) \\ u_{jk}^{G3} &\sim N(0, \sigma_{uG3}^2), v_k^{G3} \sim N(0, \sigma_{vG3}^2) \end{aligned}$$

Via Modellrestriktion wurde geprüft, ob sich die Varianzen auf Klassenebene σ_u^2 zwischen Schulen der Gruppe 1 und Gruppe 3 signifikant unterscheiden. Formal wurde folgende H0 geprüft:

$$\text{Vardiff} = \sigma_{uG1}^2 - \sigma_{uG3}^2 = 0$$

Von signifikanten Unterschieden wird gesprochen, wenn das 95%-Credibility Intervall des Parameters *Vardiff* den Wert 0 nicht einschließt.

Da es sich bei σ_{uG1}^2 und σ_{uG3}^2 um Varianzen auf einer probit-Skala handelt, wurden die Varianzen für die Ergebnisdarstellung auf Basis der Simulationsmethode von Leckie et al. (2012) in den Dissimilaritätsindex (DI; Duncan & Duncan, 1955), der eine klare und einfache Interpretation aufweist (siehe dazu auch Weber et al., 2019), transformiert. Die Simulationsmethode, die auch in Biedermann et al. (2016b) beschrieben ist, wurde in der aktuellen Anwendung auf Modellparameter aus einem Probit-Modell übertragen (Leckie et al. (2012) illustrieren die Methode für Logit-Modelle). Anzumerken ist, dass die im Beitrag dargestellten 95%-Intervalle des DI nur die durch die Simulation bedingte Unsicherheit berücksichtigen (=2,5%-Perzentil – 97,5%-Perzentil aus 1.000 simulierten

³ Dreiebenenmodelle für binären Outcomes sind in Mplus nur mit der Bayes-Schätzung und Probit-Link möglich.

⁴ Die Vorgehensweise ist in Fußnote 2 beschrieben.

Datensätzen) und nicht die Unsicherheit aufgrund der Schätzung der Varianzkomponenten. Daher fallen die Intervalle auch schmaler als beim ICC für den Sozialstatus aus.

Behandlung fehlender Werte

Für die Analysen der sozialen Segregation wurden ebenfalls die fehlenden Werte für die Indikatoren des Sozialstatus imputiert. Auch hier wurde eine multiple Imputation mit der Software Blimp 1 (Enders, Keller & Levy, 2017; Keller & Enders, 2017) durchgeführt. Im Unterschied zur oben beschriebenen Vorgehensweise wurden jedoch alle PVs der Leistungsmessungen simultan berücksichtigt und auch die Angaben der Schulleiter/innen zu den Kriterien der Klassenzusammenstellung. Insgesamt wurden 10 imputierte Datensätze erzeugt. Die Analysen zur ethnischen Segregation basieren auf einem Dreiebenen-Probit-Modell, das in Mplus nur mit einer Bayes-Schätzung geschätzt werden kann. Die Verwendung von multipel imputierten Daten ist für die Bayes-Schätzung nicht möglich. Daher wurde für die Analysen zur ethnischen Segregation keine Imputation durchgeführt. Da die Anteile der fehlenden Werte beim Migrationshintergrund sehr gering sind (BIST-Ü-M4 2013: 1,4%; BIST-Ü-D8 2016: 1,2%), ist von keiner substantiellen Verzerrung durch die Verwendung von Rohdaten auszugehen.

Ergebnisse

Nachfolgend sind die Ergebnisse dargestellt, die den Abbildungen 4.3 (Volksschule), 4.4 (AHS) und 4.5 (NMS/HS) zugrunde liegen. Für die soziale Segregation sind der ICC (Berechnung siehe oben), dessen Standardfehler (SE) und das daraus gebildete 95%-Konfidenzintervall dargestellt. Des Weiteren ist bei signifikanten Unterschieden zwischen Schulen, wo ein Zusammenstellungskriterium entscheidend (ent.) war, und jenen Schulen, wo laut Schulleitung das Kriterium gar nicht (g.n.) entscheidend war, in der Spalte Wald-Test der Testwert ($df = 1$) inklusive dem p-Wert dargestellt. Für die ethnische Segregation sind der DI und dessen 95%-Intervall, das die Unsicherheit durch die Simulation abbildet, dargestellt. Weiters sind die L2-Varianzkomponenten σ_u^2 aus dem Dreiebenen-Probit-Modell und die Standardabweichung der posterior Verteilung dargestellt. Für signifikante Unterschiede ist die Differenz $\sigma_{uG1}^2 - \sigma_{uG3}^2$ sowie deren 95%-Credibility Intervall abgebildet. Für Kriterien, die in keiner Schule bzw. in nur sehr wenigen (einstellige Anzahl an) Schulen entscheidend sind, werden keine Ergebnisse ausgewiesen (n.a.).

Tabelle 2: Unterschiede in der innerschulischen Segregation nach Zusammenstellungskriterien

		Soziale Segregation			Ethnische Segregation			
		ICC (SE)	95%-KI	Wald-Test (df = 1)	DI-Simulation DI	95%-Intervall	L2-Varianzen aus σ^2_u (Post. S.D)	Probit-Modell Diff. 95%-Cred.I
Volksschule								
der Zubringerkindergarten der Kinder	ent.	.066 (.011)	[.044; .087]	n.s.	.185	[.167; .205]	.064 (.017)	n.s.
	g.n.	.061 (.009)	[.043; .078]		.145	[.133; .156]	.044 (.01)	
örtliche/regionale Gruppen (gleiche Verkehrsmittel)	ent.	.045 (.007)	[.031; .058]	12.113; p<.001	.163	[.151; .175]	.045 (.015)	n.s.
	g.n.	.082 (.008)	[.066; .097]		.142	[.131; .154]	.048 (.01)	
Fachliche Schwerpunkte/Neigungen/Schulversuche	ent.	.117 (.013)	[.091; .142]	22.452; p<.001	.174	[.161; .187]	.068 (.015)	n.s.
	g.n.	.048 (.006)	[.036; .059]		.157	[.145; .168]	.047 (.008)	
Integrationsgesichtspunkte (von Kindern mit Behind.)	ent.	.073 (.009)	[.055; .090]	n.s.	.165	[.152; .178]	.06 (.011)	n.s.
	g.n.	.07 (.007)	[.056; .083]		.154	[.142; .167]	.047 (.01)	
Zufall / Alphabet oder ähnliches	ent.	.079 (.016)	[.047; .110]	n.s.	.088	[.081; .095]	.017 (.014)	n.s.
	g.n.	.063 (.006)	[.051; .074]		.168	[.156; .181]	.057 (.01)	
Schulreife der Schüler/innen (homogen)	ent.	.055 (.016)	[.023; .086]	n.s.	.132	[.123; .143]	.039 (.021)	n.s.
	g.n.	.069 (.006)	[.057; .080]		.167	[.155; .181]	.057 (.008)	
Schulreife der Schüler/innen (heterogen)	ent.	.054 (.007)	[.040; .067]	3.902; p<.05	.123	[.113; .133]	.034 (.09)	-.05 [-.087; -.026]
	g.n.	.078 (.01)	[.058; .097]		.210	[.195; .225]	.084 (.014)	
Herkunft/Muttersprache der Schüler/innen	ent.	.062 (.012)	[.038; .085]	n.s.	.197	[.184; .212]	.088 (.023)	.04 [.001; .090]
	g.n.	.069 (.007)	[.055; .082]		.158	[.146; .171]	.048 (.008)	
Religionsgemeinschaft/Religion der Schüler/innen	ent.	.06 (.012)	[.036; .083]	n.s.	.165	[.152; .177]	.057 (.015)	n.s.
	g.n.	.066 (.008)	[.050; .081]		.147	[.137; .158]	.041 (.009)	
Organ. Rahmenbedingungen der Schule	ent.	.069 (.026)	[.018; .119]	n.s.	.101	[.093; .108]	.021 (.019)	n.s.
	g.n.	.068 (.006)	[.056; .079]		.165	[.152; .176]	.056 (.08)	
Verhalten der Kinder/ bekannte Erz.-schwierigk.	ent.	.061 (.011)	[.039; .082]	n.s.	.145	[.135; .157]	.046 (.013)	n.s.
	g.n.	.076 (.008)	[.060; .091]		.197	[.182; .212]	.079 (.013)	
Wünsche der Eltern	ent.	.102 (.015)	[.072; .131]	1.818; p<.01	.188	[.174; .203]	.078 (.016)	n.s.
	g.n.	.044 (.009)	[.026; .061]		.159	[.139; .178]	.045 (.015)	
AHS								
Zubringerschule der Kinder	ent.	.016 (.004)	[.008; .023]	4.927; p<.05	.231	[.206; .255]	.095 (.018)	n.s.
	g.n.	.048 (.014)	[.020; .075]		.183	[.168; .197]	.075 (.028)	
örtliche/regionale Gruppen (gleiche Verkehrsmittel)	ent.	.025 (.006)	[.013; .036]	n.s.	.235	[.217; .254]	.094 (.021)	n.s.
	g.n.	.035 (.007)	[.021; .048]		.198	[.183; .214]	.093 (.017)	
Fachliche Schwerpunkte/Neigungen/Schulversuche	ent.	.02 (.005)	[.010; .029]	n.s.	.238	[.220; .257]	.115 (.029)	.049 [.003; .097]
	g.n.	.029 (.005)	[.019; .038]		.185	[.171; .200]	.066 (.014)	
Integrationsgesichtspunkte (von Kindern mit Behind.)	ent.	.031 (.012)	[.007; .054]	n.s.	.276	[.254; .298]	.162 (.063)	n.s.
	g.n.	.029 (.004)	[.021; .036]		.210	[.193; .226]	.086 (.012)	
Zufall / Alphabet oder ähnliches	ent.	.057 (.022)	[.013; .100]	n.s.	.217	[.200; .234]	.119 (.066)	n.s.
	g.n.	.026 (.005)	[.016; .035]		.220	[.203; .239]	.09 (.013)	
Homogene Schulleistungen der Schüler/innen	ent.	n.a.	n.a.	n.s.	n.a.	n.a.	n.a.	n.s.
	g.n.	n.a.	n.a.		.183	[.167; .198]	.079 (.037)	
Heterogene Schulleistungen der Schüler/innen	ent.	.049 (.024)	[.002; .096]	n.s.	.221	[.205; .238]	.092 (.013)	n.s.
	g.n.	.026 (.004)	[.018; .033]		n.a.	n.a.	n.a.	
Herkunft/Muttersprache der Schüler/innen	ent.	n.a.	n.a.	n.s.	n.a.	n.a.	n.a.	n.s.
	g.n.	n.a.	n.a.		.146	[.120; .169]	.038 (.019)	
Religionsgemeinschaft/Religion der Schüler/innen	ent.	.033 (.008)	[.017; .048]	n.s.	.238	[.211; .265]	.114 (.021)	-.076 [-.133; -.013]
	g.n.	.035 (.009)	[.017; .052]		.172	[.159; .186]	.059 (.026)	
organisatorische Rahmenbedingungen der Schule	ent.	.043 (.016)	[.011; .074]	n.s.	.215	[.200; .231]	.09 (.013)	n.s.
	g.n.	.025 (.005)	[.015; .034]		n.a.	n.a.	n.a.	
Verhalten der Kinder/ bekannte Erz.-schwierigk.	ent.	n.a.	n.a.	n.s.	n.a.	n.a.	n.a.	n.s.
	g.n.	n.a.	n.a.		.213	[.197; .231]	.091 (.026)	
Wünsche der Eltern	ent.	.024 (.008)	[.008; .039]	n.s.	.269	[.238; .300]	.147 (.034)	n.s.
	g.n.	.034 (.008)	[.018; .049]		n.a.	n.a.	n.a.	
NMS								
Zubringerschule der Kinder	ent.	.014 (.006)	[.002; .025]	4.439; p<.05	.211	[.178; .248]	.086 (.019)	n.s.
	g.n.	.032 (.006)	[.020; .043]		.243	[.217; .272]	.127 (.021)	
örtliche/regionale Gruppen (gleiche Verkehrsmittel)	ent.	.024 (.009)	[.006; .041]	n.s.	.214	[.191; .239]	.079 (.016)	n.s.
	g.n.	.03 (.006)	[.018; .041]		.196	[.174; .220]	.092 (.12)	
Fachliche Schwerpunkte/Neigungen/Schulversuche	ent.	.022 (.005)	[.012; .031]	n.s.	.279	[.248; .309]	.171 (.023)	.102 [.057; .157]
	g.n.	.024 (.005)	[.014; .033]		.216	[.194; .241]	.105 (.014)	
Integrationsgesichtspunkte (von Kindern mit Behinderungen)	ent.	.026 (.005)	[.016; .035]	n.s.	.232	[.205; .259]	.099 (.017)	n.s.
	g.n.	.026 (.007)	[.012; .039]		.134	[.117; .150]	.033 (.026)	
Zufall / Alphabet oder ähnliches	ent.	.044 (.016)	[.012; .075]	n.s.	.244	[.218; .273]	.121 (.014)	-.088 [-.132; -.012]
	g.n.	.027 (.005)	[.017; .036]		n.a.	n.a.	n.a.	
Homogene Schulleistungen der Schüler/innen	ent.	n.a.	n.a.	n.s.	n.a.	n.a.	n.a.	n.s.
	g.n.	n.a.	n.a.		.207	[.185; .230]	.093 (.017)	
Heterogene Schulleistungen der Schüler/innen	ent.	.031 (.008)	[.015; .046]	n.s.	.236	[.209; .264]	.11 (.01)	n.s.
	g.n.	.024 (.005)	[.014; .033]		n.a.	n.a.	n.a.	
Herkunft/Muttersprache der Schüler/innen	ent.	n.a.	n.a.	n.s.	n.a.	n.a.	n.a.	n.s.
	g.n.	n.a.	n.a.		.184	[.162; .206]	.077 (.036)	
Religionsgemeinschaft/Religion der Schüler/innen	ent.	.015 (.009)	[-.002; .032]	n.s.	.238	[.208; .267]	.11 (.012)	n.s.
	g.n.	.03 (.015)	[.001; .059]		.247	[.220; .276]	.143 (.043)	
Organ. Rahmenbedingungen der Schule	ent.	.013 (.009)	[-.005; .030]	n.s.	.217	[.192; .243]	.096 (.01)	n.s.
	g.n.	.028 (.004)	[.0201; .035]		.189	[.169; .211]	.083 (.035)	
Verhalten der Kinder/ bekannte Erz.-schwierigk.	ent.	.001 (.013)	[-.024; .026]	n.s.	.241	[.215; .266]	.116 (.015)	n.s.
	g.n.	.026 (.004)	[.018; .033]		n.a.	n.a.	n.a.	

Wünsche der Eltern	ent.	.007 (.007)	[-.006; .020]	n.s.	.229	[.200; .258]	.127 (.031)	
	g.n.	.024 (.006)	[.0122; .035]		.237	[.197; .275]	.101 (.021)	n.s.

C Verteilung der Lehrkraftqualität innerhalb von Schulen

In Abschnitt 2.4.2 wurde auf Basis der Talis-2008-Daten (OECD, 2010) der Frage nachgegangen, ob innerhalb von Schulen Zusammenhänge zwischen Merkmalen der Klassenzusammensetzung und Indikatoren der Lehrkraftqualität bestehen.

Analysemodelle

Als Indikatoren der Lehrkraftqualität (*LeQual*) werden die Erfahrung und die fachliche Qualifikation verwendet (siehe auch Allen & Sims, 2018). Die Erfahrung hat einen Wertebereich von 1 (=1. Dienstjahr) bis 7 (=mehr als 20 Dienstjahre). Die fachliche Qualifikation ist eine dichotome Variable, die abbildet, ob das Fach, das in der zu wählenden Referenzklasse⁵ unterrichtet wird, auch Teil der Ausbildung war (0 = ja, 1 = nein). Als Indikatoren der Zusammensetzung der Referenzklasse (*klazu*) werden drei 5-stufige ordinale Variablen verwendet: (1) Anteil Schüler/innen mit nichtdeutscher Muttersprache (%ndM) und (2) Anteil Schüler/innen, bei denen zumindest ein Elternteil einen akademischen Abschluss, eine Meisterprüfung oder einen anderen Abschluss auf ISCED5-Niveau aufweist (%E_ISCED5). Beide Aspekte wurden von den Lehrkräften auf einer 5-stufigen Skala (1 = weniger als 10 %; 2= 10% oder mehr, aber weniger als 20%; 3 = mehr als 20%, aber weniger als 40 %; 4 = mehr als 40 %, aber weniger als 60%; 5 = 60 % oder mehr) eingeschätzt. (3) Eingeschätzte Leistung der Referenzklasse im Vergleich zu anderen Klassen der gleichen Schulstufe an der jeweiligen Schule mit den Ausprägungen 1 = Viel niedriger als der Durchschnitt, 2 = ein wenig niedriger als der Durchschnitt, 3 = durchschnittlich, 4= ein wenig höher als der Durchschnitt und 5 = viel höher als der Durchschnitt. Analog zu Allen und Sims (2018) wurden Zweiebenen-Logit-Modelle für ordinale Variablen (u.a. Raudenbush & Bryk, 2002, S. 317ff.) mit den Indikatoren der Klassenzusammensetzung (*klazu*) als abhängige Variablen und den Indikatoren der Lehrkraftqualität (*LeQual*) als unabhängigen Variablen gerechnet. Folgendes Modell liegt den Analysen zugrunde:

$$klazu_{ij} \sim multinomial(\pi_{ij}^{(1)}, \pi_{ij}^{(2)}, \pi_{ij}^{(3)}, \pi_{ij}^{(4)})$$

$$\log\left(\frac{\pi_{ij}^{(1)}}{\pi_{ij}^{(2,3,4,5)}}\right) = \beta_0^{(1)} + \beta_1 LeQual_{ij} + u_j$$

$$\log\left(\frac{\pi_{ij}^{(1,2)}}{\pi_{ij}^{(3,4,5)}}\right) = \beta_0^{(2)} + \beta_1 LeQual_{ij} + u_j$$

$$\log\left(\frac{\pi_{ij}^{(1,2,3)}}{\pi_{ij}^{(4,5)}}\right) = \beta_0^{(3)} + \beta_1 LeQual_{ij} + u_j$$

$$\log\left(\frac{\pi_{ij}^{(1,2,3,4)}}{\pi_{ij}^{(5)}}\right) = \beta_0^{(4)} + \beta_1 LeQual_{ij} + u_j$$

$$\mu_j \sim N(0, \sigma_u^2)$$

$\pi_{ij}^{(1)}$ ist die Wahrscheinlichkeit für die Kategorie 1 der Klassenzusammensetzung (weniger als 10% bzw. viel niedriger als der Durchschnitt), $\pi_{ij}^{(2)}$ ist die Wahrscheinlichkeit der Kategorie 2 usw... $\pi_{ij}^{(1,2,3)}$ ist die kumulative Wahrscheinlichkeit der Kategorien 1 bis 3. Kategorie 5 stellt die Referenzkategorie dar ($\pi_{ij}^{(5)}$ ist redundant, weil $1 - \pi_{ij}^{(1,2,3,4)} = \pi_{ij}^{(5)}$). Die Modellgleichungen beschreiben die log odds von Kategorie 1 vs. Kategorien größer 1 durch

⁵ Die Wahl der Referenzklasse erfolgt wie im folgenden Beispiel: „Die Klasse, auf die Sie sich beziehen sollen, ist die erste Klasse der Sekundarstufe 1, die Sie gewöhnlich dienstags nach 11 Uhr unterrichten. Wenn Sie am Dienstag keine Klasse mehr unterrichten, dann beziehen Sie sich auf die Klasse, die Sie als erstes am folgenden Tag, an dem Sie Unterricht haben, unterrichten.“

einen spezifischen Schwellenwertparameter $\beta_0^{(1)}$ und einen gemeinsamen Regressionskoeffizienten β_1 . Auf Basis dieser Parameter ($\beta_0^{(1)}, \beta_0^{(2)}, \beta_0^{(3)}, \beta_0^{(4)}$ und β_1) wurden für die Darstellung im Beitrag Wahrscheinlichkeiten für die einzelnen Kategorien bzw. kumulative Wahrscheinlichkeiten (z. B. $\pi_{ij}^{(1,2,3)}$ = Wahrscheinlichkeit, dass die abhängige Variable eine Ausprägung < 4 annimmt) berechnet (siehe dazu Muthén & Muthén, 1998-2017, S. 556). Da für die Analysen der innerschulische Effekt (also der Vergleich zu anderen Lehrkräften der Schule) im Zentrum des Interesses steht, wurde LeQual als jeweiliger Schulmittelwert zentriert (Enders & Tofighi, 2007). β_1 gibt somit den innerschulischen Zusammenhang zwischen Lehrkraftqualität und Klassenzusammensetzung wieder. Für alle Analysen wurde die GewichtungsvARIABLE TCHWGT verwendet (Dumais & LaRoche, o.J., S. 23). Es wurde keine multiple Imputation durchgeführt. Die fehlenden Werte reichen von rund 1% bis zu ca. 10% (Anteil ISCED-5). Die vorliegenden Analysen sollen eine erste Befundlage für Österreich mit vergleichsweise alten Daten bieten. Für zukünftige Analysen (z.B. mit TALIS 2018 Daten) sollte daher eine adäquate Behandlung fehlender Werte erfolgen.

Ergebnisse

Nachfolgend sind die Ergebnisse der Analysen dargestellt. Es zeigt sich beispielsweise ein negativer Zusammenhang zwischen Erfahrung und Anteil der Schüler/innen mit Eltern, die zumindest eine Ausbildung auf ISCED-5 Niveau abgeschlossen haben. D.h. je höher die Erfahrung, desto geringer der ISCED-5-Anteil.

Tabelle 3: Klassenzusammensetzung und Lehrkräftemerkmale

	HS			AHS		
	% ndM β_1 (SE), OR	% E_ISCED5 β_1 (SE), OR	Leistung β_1 (SE), OR	% ndM β_1 (SE), OR	% E_ISCED5 β_1 (SE), OR	Leistung β_1 (SE), OR
Erfahrung	.016 (.039), 1.016	-.123*** (.027), .885	-.001 (.030), .999	.037 (.046) 1.038	-.147*** (.031), .863	-.016 (.032) .984
Keine fachliche Qualifikation	.323* (.131), 1.382	.027 (.120), 1.027	-.481***, (.108) .618	- ^a	- ^a	- ^a

Anmerkungen: OR = Odds Ratio, SE = Standardfehler. %ndM = Anteil Schüler/innen mit nichtdeutscher Muttersprache, % E_ISCED5 = Anteil Schüler/innen bei denen zumindest ein Elternteil einen akademischen Abschluss, eine Meisterprüfung oder einen anderen Abschluss auf ISCED5 aufweist. Leistung = Eingeschätzte Leistung der Referenzklasse im Vergleich zu anderen Klassen der gleichen Schulstufe an der jeweiligen Schule. Ausprägungen für %ndM und % E_ISCED5: weniger als 10 %; 10% oder mehr, aber weniger als 20%; mehr als 20%, aber weniger als 40 %; mehr als 40 %, aber weniger als 60%; 60 % oder mehr. Ausprägungen für Leistung: Viel niedriger als der Durchschnitt, ein wenig niedriger als der Durchschnitt, durchschnittlich, ein wenig höher als der Durchschnitt, viel höher als der Durchschnitt. Gerechnet wurde jeweils ein Zweiebenen-Logit-Modell für ordinale abhängige Variablen. Die unabhängigen Variablen auf Lehrkräfteebene (Erfahrung und fehlende fachliche Qualifikation) wurden am Schulmittelwert zentriert. ^a ... Fachfremder Unterricht ist unter AHS-Lehrkräften sehr selten (ca. 2,8%). Daher wurden hier keine Zusammenhänge berechnet.

Literatur

- Allen, R. & Sims, S. (2018). Do pupils from low-income families get low-quality teachers? Indirect evidence from English schools. *Oxford Review of Education*, 44 (4), 441–458. <http://doi.org/10.1080/03054985.2017.1421152>
- Asparouhov, T. & Muthén, B. (2010). Bayesian Analysis Using Mplus: Technical Implementation. Mplus Technical Report.
- Biedermann, H., Weber, C., Herzog-Punzenberger, B. & Nagel, A. (2016a). Auf die Mitschüler/innen kommt es an? Schulische Segregation – Effekte der Schul- und Klassenzusammensetzung in der Primarstufe und der Sekundarstufe I. In M. Bruneforth, F. Eder, K. Krainer, C. Schreiner, A. Seel & C. Spiel (Hrsg.), *Nationaler Bildungsbericht Österreich 2015, Band 2: Fokussierte Analysen bildungspolitischer Schwerpunktthemen* (S. 133–174). Graz: Leykam. <http://dx.doi.org/10.17888/nbb2015-2-4>
- Biedermann, H., Weber, C., Herzog-Punzenberger, B. & Nagel, A. (2016b). Schulische Segregation: Technische Dokumentation und Detailergebnisse. Onlinedokument zum Kapitel „Auf die Mitschüler/innen kommt es an? Schulische Segregation – Effekte der Schul- und Klassenzusammensetzung in der Primarstufe und der Sekundarstufe I“ im Band 2 des Nationalen Bildungsberichts Österreich 2015. <http://dx.doi.org/10.17888/nbb2015-2-4-1>
- Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens (BIFIE). (2015). *Standardüberprüfung 2013 Mathematik, 4. Schulstufe* (BIST-Ü-M4 2013). Unveröffentlichter Datensatz. Salzburg: BIFIE.
- Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens (BIFIE). (2017a). *Standardüberprüfung 2015 Deutsch, 4. Schulstufe* (BIST-Ü-D4 2015). Unveröffentlichter Datensatz. Salzburg: BIFIE.
- Bundesinstitut für Bildungsforschung, Innovation & Entwicklung des österreichischen Schulwesens (BIFIE). (2017b). *Standardüberprüfung 2016 Deutsch, 8. Schulstufe* (BIST-Ü-D8 2016). Unveröffentlichter Datensatz. Salzburg: BIFIE.
- Cohen, J. (1992). A power primer. *Psychological bulletin*, 112 (1), 155–159.
- Duncan, O. D. & Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review*, 20 (2), 210–217.
- Dumais, J. & LaRoche, S. (o.J.). Weights and complex sample variance estimation. In C. Carsten (Eds.), *TALIS: User guide for the international database* (S. 23-26). Paris: OECD.
- Eder F., Altrichter, H., Hofmann, F. & Weber, C. (Hrsg.). (2015). Evaluation der Neuen Mittelschule (NMS). Befunde aus den Anfangskohorten. Forschungsbericht. Graz: Leykam. Zugriff am 27.10.2015 unter https://www.bifie.at/system/files/dl/eval_forschungsbericht.pdf
- Enders, C. K., Keller, B. T., & Levy, R. (2017). A chained equations imputation approach for multilevel data with categorical and continuous variables. *Psychological Methods*, Advance online publication. <http://dx.doi.org/10.1037/met0000148>.
- Enders, C. K. & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological methods*, 12 (2), 121.
- Freunberger, R., Robitzsch, A. & Pham, G. (2014). Hintergrundvariablen und spezielle Analysen in der BIST-Ü-M4 2013. Technischer Bericht. Verfügbar am 08.09.2015 unter <https://www.bifie.at/node/2765>
- Kaplan, D., & Su, D. (2018). On imputation for planned missing data in context questionnaires using plausible values: a comparison of three designs. *Large-scale Assessments in Education*, 6(1), 6.
- Keller, B. T., & Enders, C. K. (2017). *Blimp User's Manual (Version 1.0)*. Los Angeles, CA.
- Leckie, G., Pillinger, R., Jones, K. & Goldstein, H. (2012). Multilevel modeling of social segregation. *Journal of Educational and Behavioral Statistics*, 37 (1), 3–30.
- Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Morin, A. J., Abduljabbar, A. S. & Köller, O. (2012). Classroom climate and contextual effects: Conceptual and methodological issues in the evaluation of group-level effects. *Educational Psychologist*, 47 (2), 106–124.
- Marsh, H. W., Kuyper, H., Morin, A. J., Parker, P. D., & Seaton, M. (2014). Big-fish-little-pond social comparison and local dominance effects: Integrating new statistical models, methodology, design, theory and substantive implications. *Learning and Instruction*, 33, 50-66.
- Muthén, L. K. and Muthén, B. O. (1998–2017). *Mplus User's Guide*. Eight Edition. Los Angeles, CA: Muthén & Muthén.

Organisation for Economic Cooperation and Development (OECD). (2010). TALIS 2008 technical report. Paris: OECD Publishing. Verfügbar unter <http://www.oecd.org/education/school/44978960.pdf>

Raudenbush, S.W. & Bryk, A.S. (2002). Hierarchical Linear Models. Applications and Data Analysis Methods. Second Edition. Thousand Oaks: Sage.

Rubin, D. B. (2003). Nested multiple imputation of NMES via partially incompatible MCMC. *Statistica Neerlandica*, 57(1), 3-18.

Weber, C., Moosbrugger, R., Hasenberger, K., Altrichter, H. & Schrodt, H. (2019). Wer unterrichtet wen? Die Zusammensetzung von Klassen und Schulen und die Zuteilung von Lehrkräften. In S. Breit, F. Eder, K. Krainer, C. Schreiner, A. Seel & C. Spiel (Hrsg.), *Nationaler Bildungsbericht Österreich 2018, Band 2: Fokussierte Analysen und Zukunftsperspektiven für das Bildungswesen* (S. 143–182). Graz: Leykam. <http://doi.org/10.17888/nbb2018-2-4>