

Klassenheterogenität und Leistungsentwicklung in der Primarstufe

Die Anwesenheit von Schüler*innen mit Sonderpädagogischem Förderbedarf und die Leistungsentwicklung ihrer Mitschüler*innen

Mathias Krammer

Abstract

In diesem Beitrag werden die Auswirkungen inklusiver Beschulungsmodelle auf die Leistungsentwicklung in Deutsch/Lesen und Mathematik von Schüler*innen ohne Sonderpädagogischen Förderbedarf untersucht. Dabei liegt der Fokus zum einen auf dem Auffinden und der Stärke von möglichen (linearen) Kompositionseffekten hinsichtlich der Leistungsentwicklung, basierend auf der Anwesenheit von Schüler*innen mit Sonderpädagogischem Förderbedarf im Klassenverband. Zum anderen wurde auch untersucht, ob inklusive Klassenzusammensetzungen existieren, die sich förderlich auf die Leistungsentwicklung aller Schüler*innen einer Klasse auswirken.

Dazu wurden mit den Bildungsstandarddaten für die 4. Schulstufe Mathematik (2013) und Deutsch/Leseverständnis (2015) eine Reihe von Mehrebenenanalysen (MLM) berechnet. Die Analyse der Leistungsdaten dieser beiden Vollerhebungen ($N_{\text{Gesamt}} = \text{ca. } 160.000$) zeigt von der Effektstärke her nur sehr schwach ausgeprägte Effekte, die je nach Modell sowohl positive als auch negative Werte aufweisen. Insgesamt ist daher hinsichtlich der Anwesenheit von Kindern mit Sonderpädagogischem Förderbedarf in der Primarstufe – im Gegensatz zu anderen Kompositionsmerkmalen wie sozio-ökonomischer Status oder Selbstkonzept – nur von einem sehr geringen, nicht-linearen Effekt auszugehen (*Cohens d*: -0,16 bis 0,1).



Einleitung

Bei diesem Beitrag handelt es sich um eine Zusammenfassung zweier bereits publizierter Artikel, welche in den englischsprachigen Zeitschriften „International Journal of Inclusive Education“ und „Educational Studies“ bereits in ausführlicher Form erschienen sind (Krammer et al., 2019; Krammer et al., 2021). Diese Publikationen beschäftigen sich mit den Auswirkungen inklusiver Beschulungsmodelle auf die Leistungsentwicklung innerhalb der Primarstufe. Nachdem bereits eine Reihe von Studien sich mit der positiven Wirkung inklusiver Beschulungsmodelle auf die Leistungsentwicklung von Schüler*innen mit Sonderpädagogischem Förderbedarf (SPF) auseinandersetzt und die positive Wirkung dieser Art von Beschulung auf diese Schülergruppe international belegt werden konnte (Dessementet et al., 2012; Haerberlin et al., 1991), untersuchen diese empirischen Arbeiten die Auswirkungen inklusiver Beschulungsmodelle auf Schüler*innen ohne Sonderpädagogischen Förderbedarf.

Während die Auswirkungen von inklusiven Beschulungsmodellen auf die schulische Leistungserreichung von Schüler*innen mit SPF bereits gut erforscht und dokumentiert sind, bleibt der Forschungsstand hinsichtlich der Auswirkung inklusiver Beschulungsmodelle auf die Schulleistungen von Schüler*innen ohne SPF überschaubar. In der entsprechenden Literatur werden dabei vor allem zwei dichotome Erklärungsansätze verfolgt. Zum einen wird davon ausgegangen, dass inklusive Beschulungsmodelle eine positive Wirkung auf die schulische Leistungserreichung auch von Schüler*innen ohne SPF aufweisen (Ruijs, 2017). Dies wird vor allem damit argumentiert, dass die zusätzlichen Ressourcen, die in inklusiven Schulklassen vorhanden sind (z. B. zusätzliche Lehrkraft), oft in Verbindung mit einer verstärkt betriebenen Individualisierung und Differenzierung des Unterrichts, allen Schüler*innen einer Klasse zugutekommt. Allerdings muss hier auch kritisch angemerkt werden, dass gerade die zusätzlichen Ressourcen, die in einer inklusiven Schulklasse vorhanden sind, ja auch gerade von den Schüler*innen mit SPF benötigt und damit auch ein Mehr an Aufmerksamkeit und Ressourcen der Lehrperson beansprucht werden. Damit bleibt dann weniger Zeit und Aufmerksamkeit der Lehrperson für die anderen Schüler*innen in der Klasse, was wiederum zu negativen Effekten auf die schulische Leistungserreichung dieser Schüler*innen führen kann (Hienonen et al., 2018; Kristoffersen et al., 2015).

In diesem Zusammenhang ist es daher auch nicht verwunderlich, dass sich auch bei empirischer Betrachtung ein etwas uneinheitliches Bild ergibt, bei dem jedoch in letzter Zeit die positiven Befunde zu inklusiven Beschulungsmodellen überwiegen. So zeigten beispielsweise Szumski, Smogorzewska und Karwowski (2017) in einer fast fünf Millionen Schüler*innen umfassenden Metaanalyse, dass inklusive Beschulungsmodelle im Durchschnitt einen schwach positiven Effekt auf die schulische Leistungserreichung auch der Schüler*innen ohne SPF aufweisen. Ähnliche Ergebnisse sind zudem auch bei Kalambouka et al. (2007) und bei einer Reihe von weiteren Studien zu finden (Cosier et al.,

2013; Hanushek et al., 2002). Allerdings muss hier auch darauf hingewiesen werden, dass gerade die Anwesenheit von Schüler*innen mit einem SPF in Bereich Verhalten auch einen negativen Effekt auf die schulische Leistungserreichung hervorrufen kann (Krisoffersen et al., 2015). Schwach negative Effekte auf die Leistungserreichung der Schüler*innen ohne Special Educational Needs (SEN) wies zudem Hienonen et al. (2018) für eine verhältnismäßig große finnische Stichprobe nach. Zusammenfassend bleibt jedoch festzuhalten, dass bei den meisten Studien eine relativ geringe Effektstärke auf die Leistungserreichung (positiv wie negativ) festgestellt wurde (Krammer et al., 2021).

Allerdings muss auch darauf hingewiesen werden, dass die meisten Untersuchungen sich allein auf die Anwesenheit von Schüler*innen mit SPF konzentrieren. Weitere Kompositionseffekte wurden bei einem Großteil der Studien nicht berücksichtigt. Dies ist v. a. dahingehend bedeutsam, da gerade Kompositionseffekte aufgrund des sozioökonomischen Status in einer Schulklasse deutliche Auswirkungen auf die schulische Leistungserreichung sowohl in Deutsch als auch in Mathematik haben (Niklas & Schneider, 2013; Philipp, 2011; Rodriguez-Brown, 2011, Stubbe et al., 2016). So zeigten etwa Groeneveld und Knigge (2015), dass bis zu 20 % der Varianz im Leseverständnis bei Sechstklässlern durch die sozioökonomische Zusammensetzung der Klassen erklärt werden kann. Ähnliche Auswirkungen wurden auch für Mathematik nachgewiesen (Sirin, 2005). Stubbe et al. (2016) zeigten bei der TIMMS Studie für Deutschland, dass ebenfalls zwischen 13 und 18 % der Varianz der Mathematikleistung durch sozioökonomische Kompositionseffekte erklärt werden können. Ein sehr ähnliches Bild ergibt sich in dieser Hinsicht auch für Österreich. So belegten Biedermann et al. (2016) sowohl für die Primar- als auch Sekundarstufe, dass der sozioökonomische Status ein bedeutender Prädiktor sowohl für Mathematik als auch für Deutsch darstellt. Da zudem ein Zusammenhang zwischen der Vergabe von SPFs und dem sozioökonomischen Status der Schüler*innen besteht, sollten entsprechende Untersuchungen auch diese Art von Kompositionseffekten in den Berechnungen berücksichtigen, was bisher allerdings bei Weitem nicht in allen Untersuchungen der Fall war (Shifrer et al., 2011).

Daneben spielen eine Vielzahl an weiteren individuellen als auch schulischen Einflussfaktoren (Geschlecht, Qualität des Unterrichts, kognitive Leistungsfähigkeit, Selbstkonzept etc.) eine Rolle bei der schulischen Leistung. Sowohl für Lesen als auch für Mathematik wurden dabei entsprechende Zusammenhänge zwischen dem Geschlecht und Schulleistung nachgewiesen. Für das Lesen zeigen eine Reihe von internationalen Large-scale-Studien, dass Mädchen im Durchschnitt deutlich besser abschneiden als dies Jungen tun (Reilly et al., 2018). Im Gegenstand Mathematik ist die Richtung dieses Zusammenhangs diametral, zumindest für Österreich. Für PISA 2015 erreichen Jungen im Schnitt 27 Punkte mehr als Mädchen; ein Ergebnis, das auch über den Zeitverlauf vergleichsweise stabil ist (Suchan & Breit, 2016). Des Weiteren spielt auch der Kindergartenbesuch eine Rolle für das Leseverständnis. Dabei zeigt sich länderübergreifend, dass die Dauer des Kindergartenbesuchs einen positiven Zusammenhang mit dem Leseverständnis auf-

weist: Je länger Schüler*innen den Kindergarten besuchen, desto besser ist ihr späteres Leseverständnis (Anders, 2013). Schließlich konnten bezüglich des Selbstkonzepts auch hier positive Zusammenhänge zwischen dem fachlichen Selbstkonzept und der Schulleistung (neben anderen Faktoren wie etwa der Motivation) festgestellt werden (Wagner et al., 2008).

Forschungsfragen

Die vorliegende Arbeit untersucht für die österreichische Primarstufe, welchen Einfluss die Anwesenheit von Schüler*innen mit SPF auf das Abschneiden der übrigen Schüler*innen in der Klasse bei den Bildungsstandardüberprüfungen hat. Gleichzeitig wurden der sozioökonomische Status (sowohl bei Leseverständnis als auch bei Mathematik) und das Selbstkonzept in der Klasse (nur für die Daten der BIST-Deutsch) in die Analysen miteinbezogen. Dazu wurden folgende Forschungsfragen formuliert:

1. Beeinflusst die Anzahl/Anwesenheit von Schüler*innen mit diagnostiziertem SPF das Abschneiden der übrigen Schüler*innen im Klassenverbund bei der BIST-Überprüfung für Deutsch (Leseverständnis) und Mathematik?
 - a) Existiert in diesem Zusammenhang ein möglicher positiver/negativer linearer Zusammenhang?
 - b) Unterscheiden sich mögliche Effekte bei leistungsschwachen Schüler*innen ohne SPF von jenen bei leistungsstarken Schüler*innen?
2. Wie stark sind mögliche Effekte auf die Leistungserreichung vor allem auch im Vergleich mit bekannten Kompositionseffekten wie etwa dem sozio-ökonomischen Status (Leseverständnis und Mathematik) oder dem Selbstkonzept (nur für das Leseverständnis).

Daneben wurde noch eine Reihe von Individualvariablen (Geschlecht, Alter und für das Leseverständnis zusätzlich die Dauer des Kindergartenbesuchs und das Selbstkonzept) in den Regressionsmodellen berücksichtigt, die einen bekannten Zusammenhang mit den Leistungsvariablen aufweisen.

Methoden

Um diese Fragen zu beantworten, wurde eine Reihe von Mehrebenenmodellen, basierend auf den Daten der Bildungsstandardüberprüfungen Mathematik 2013 (M4-2013) und Deutsch 2015 (D4-2015) für die vierten Schulstufen, berechnet. Alle Berechnungen wurden mittels des R-Packages „BIFIESurvey“ in R-Studio durchgeführt. Dabei wurden die Berechnungen zuerst an einem vom IQS zur Verfügung gestellten Subsample mit ca. 4000 Schüler*innen durchgeführt. Anschließend wurde der dabei generierte Code auf die Gesamtdatensätze angewandt. Es handelt sich daher um eine Sekundärdatenanalyse bereits vorhandener Daten. Aus diesem Grund unterschieden sich die Modelle für Deutsch und Mathematik auch etwas voneinander (unterschiedliche Prädiktoren), da bei beiden Standardüberprüfungen nicht die gleichen Variablen erhoben wurden.

Stichproben

Im Folgenden werden die Stichprobencharakteristika jeweils getrennt für Deutsch und Mathematik dargestellt. Resampling-Methoden und Multiple Imputations wurden zur Schätzung der Sampling-Varianz und zur Schätzung von fehlenden Werten herangezogen. Für eine detaillierte Beschreibung der verwendeten Methoden siehe auch BIFIE (2013).

Mathematik

Insgesamt nahmen 73.655 Schüler*innen an der BIST Überprüfung Mathematik für die vierten Schulstufen im Jahr 2013 teil. Die in dieser Untersuchung verwendeten Daten stellen daher eine Gesamterhebung aller Viertklässler*innen für das Jahr 2013 dar, unter Ausschluss der außerordentlichen Schüler*innen und jenen mit diagnostiziertem SPF. Die getesteten Schüler*innen besuchten insgesamt 4.904 Klassen in 3.048 Schulen. Ca. 51 % der Schüler*innen waren weiblich und im Durchschnitt waren die teilnehmenden Schüler*innen 10,34 Jahre alt (SD: 0,45).

Deutsch

Bei der BIST-Testung für Deutsch der vierten Klassen 2015 nahmen insgesamt 75.341 Schüler*innen teil. Diese Schüler*innen besuchten insgesamt 4.927 Klassen in insgesamt 2.994 Schulen. Im Schnitt waren die Schüler*innen 10,36 Jahre alt (SD: 0,46). 49,4 % der getesteten Schüler*innen waren weiblich und ca. 81 % gaben Deutsch als Muttersprache an. Analog zur BIST-Testung für Mathematik stellen die verwendeten Daten eine Gesamterhebung aller Viertklässler*innen für das Jahr 2015 dar, ebenfalls unter Ausschluss der außerordentlichen Schüler*innen und jenen mit diagnostiziertem SPF.

Variablen

Als abhängige Variablen wurden jeweils das Leseverständnis für den Deutsch-Datensatz und die Standard-Scores in Mathematik für den Mathematikdatensatz herangezogen. Beide abhängige Variablen weisen eine Normalverteilung auf (Mathematik: $M=533,17$, $SD=99,89$; Deutsch: $M=523$, $SD=100,48$). Der Interquartilsabstand, jener Wert, in dem 50 % der getesteten Schüler*innen liegen, beträgt für Deutsch 454–594 und für Mathematik 464–603.

Neben dem Geschlecht und der Erstsprache der Schüler*innen wurde auch eine Reihe sozialer Hintergrundvariablen, wie etwa Bücher im Haushalt, Ausbildung der Eltern bzw. der Highest International Socio-Economic Index of Occupational Status (HISEI) des Elternhauses sowohl für Deutsch als auch Mathematik, einbezogen. Für Deutsch wurde dazu ein Sozialstatus-Index gebildet, der sich aus dem HISEI, den Büchern im Haushalt und der Ausbildung der Eltern zusammensetzt, während in den Mathematikmodellen diese Variablen getrennt berücksichtigt wurden.¹ Zudem wurde für den Deutschdatensatz auch die Dauer des Kindergartenbesuchs als Individualvariable herangezogen.

Neben diesen Variablen zum individuellen Hintergrund wurden auch spezifische Variablen für den jeweiligen Gegenstand in den dazugehörigen Modellen verwendet. In den Modellen für das Leseverständnis sind das die Dauer des Kindergartenbesuchs und das Selbstkonzept im Gegenstand Deutsch. Das Selbstkonzept wurde mittels eines Likert-skalierten Instruments erhoben, aus welchem anschließend ein Mittelwert gebildet wurde ($M=3,1$; $SD=0,8$, siehe auch Kramer et al., 2021). Zudem wurde für die Berechnungsmodelle zur Deutschstandardüberprüfung auch die Klassengröße (Anzahl der Schüler*innen einer Klasse) berücksichtigt.

Für den Mathematikdatensatz waren keine Angaben zum Selbstkonzept der getesteten Kinder vorhanden. Stattdessen konnten jedoch in den Rechenmodellen Angaben zur Anzahl der Schüler*innen mit einer anderen Erstsprache als Deutsch verwendet werden.

Die Anzahl der Schüler*innen mit SPF wurde als unabhängige Variable auf Klassenebene genutzt. Von den 4.927 an der Standardtestung für Deutsch teilnehmenden Klassen hatten 76,6 % keine Schüler*innen mit SPF, 13 % ein Kind mit diagnostiziertem SPF, 4 % zwei Schüler*innen mit SPF und ca. 5 % drei oder mehr. Relativ ähnliche Zahlen ergeben sich für die Standardtestung in Mathematik. In 75 % der 4.904 getesteten Klassen gab es keine Schüler*innen mit diagnostiziertem SPF, in 12 % der Klassen ein Kind mit SPF, in 5 % zwei Kinder und in etwas mehr als 5 % drei oder mehr.

1 Bildung Sozial Status Index: $S = 1/6 * (z(\text{HISEI}_{\text{Eltern}}) + z(\text{HISEI}_{\text{Schüler*innen}}) + z(\text{BBücher}_{\text{Eltern}}) + z(\text{Bücher}_{\text{Schüler*innen}})) + 1/3 * z(\text{Ausbildung}_{\text{Eltern}})$

Modelle

Um den Einfluss der Klassenebene adäquat zu berücksichtigen, wurde eine Reihe von Mehrebenenmodellen jeweils getrennt für beide Datensätze berechnet. Die mathematische Formulierung für das allgemeine Modell (1) und das Leermodell ist sowohl für Deutsch als auch für Mathematik ident. Wobei y_{ij} die abhängige Variable (Standardscores für Deutsch oder Mathematik), X_{ij} die festen Effekte und Z_{ij} die Zufallseffekte darstellt.

$$y_{ij} = X_{ij}\gamma + Z_{ij}\mu_j + \epsilon_{ij} \quad (1)$$

$$y_{ij} = X_0\gamma + Z_0\mu_j + \epsilon_{ij} \quad (\text{Leermodell})$$

Die mathematische Formulierung von Individual- und Klassenebene unterscheidet sich aufgrund unterschiedlicher Variablen voneinander. Für Mathematik wurden folgende Modelle formuliert:

$$y_{ij} = X_{i(\text{Geschlecht, Alter, Bildung Eltern, Bücher, HISEI})j} \gamma + Z_0\mu_j + \epsilon_{ij} \quad (2)$$

$$y_{ij} = X_{ij(\text{Anzahl Schüler SPF, Anteil Erstsprache})} \gamma + Z_0\mu_j + \epsilon_{ij} \quad (3)$$

$$y_{ij} = X_{i(\text{Geschlecht, Alter, Bildung Eltern})j(\text{Anzahl Schüler SPF, Anteil Erstsprache})} \gamma + Z_0\mu_j + \epsilon_{ij} \quad (4)$$

Die Modelle drei und vier stellen jeweils das Individualmodell und das Klassenmodell dar, Modell fünf schließlich das beide Ebenen umfassende Gesamtmodell. Dieses Gesamtmodell wurde für den Gesamtdatensatz (4) und auch jeweils auf zwei nach der Leistungsvariable gefilterte Datensätze, je für das untere und obere Leistungsquartil, berechnet, wobei zudem auch die Anzahl der Schüler*innen mit diagnostiziertem SPF dichotomisiert wurde (Modell 5 ein*e Schüler*in mit SPF in der Klasse, Modell 6 zwei bzw. drei Schüler*innen mit SPF in der Klasse, Modell 7 vier Schüler*innen mit SPF in der Klasse).

Für den Deutschdatensatz wurden neben dem Alter und dem Geschlecht eine aggregierte Sozialstatus-Variable und das Selbstkonzept auf Individualebene verwendet (Modell 8). Im Gegensatz zum Mathematikdatensatz wurden Individual- und Klassenebene gleich zu einem Gesamtmodell zusammengefasst (9)

$$y_{ij} = X_{i(\text{Geschlecht, Alter, Sozialstatus, Selbstkonzept, Erstsprache})j} \gamma + Z_0\mu_j + \epsilon_{ij} \quad (8)$$

$$y_{ij} = X_{i(\text{Geschlecht, Alter, Sozialstatus, Selbstkonzept, Erstsprache})j(\text{Anzahl Schüler SPF, Klassengröße, Sozialstatus Klasse, Selbstkonzept Klasse})} \gamma + Z_0\mu_j + \epsilon_{ij} \quad (9)$$

Analog zum Vorgehen beim Mathematikdatensatz wurden die jeweiligen Berechnungen des Gesamtmodells ebenfalls auf nach der Leistungsvariable gefilterten Datensätzen durchgeführt (Modelle 10 und 11 für unteres und oberes Leistungsquartil, ohne Dichotomisierung der Schüler*innen mit SPF).

Ergebnisse

In Abbildung 1 sind die Verteilung der Mittelwerte der Standardüberprüfung in Mathematik, jeweils nach Anzahl der Schüler*innen mit SPF in der Klasse, deskriptiv dargestellt.

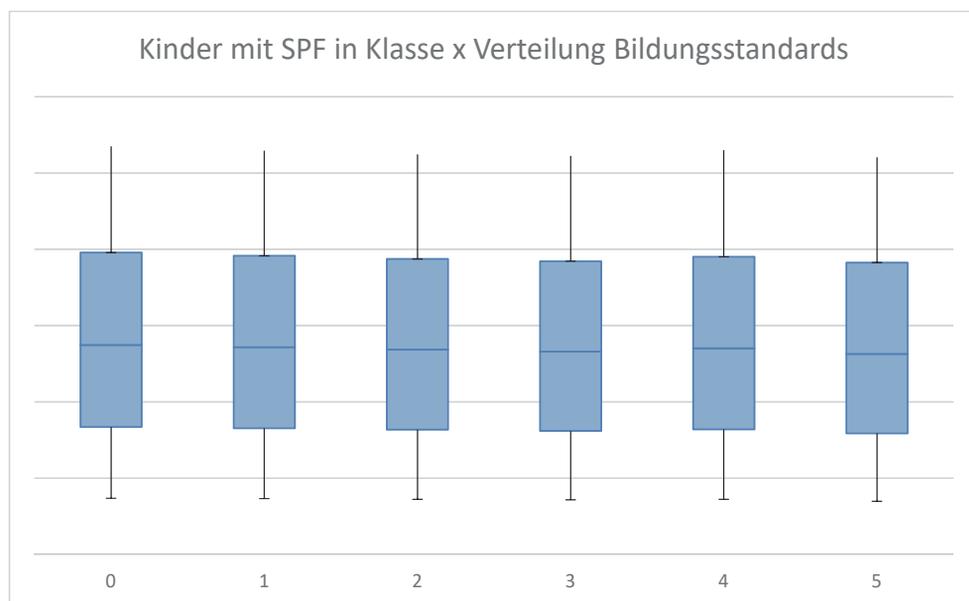


Abb. 1: Anzahl der Schüler*innen mit SPF in der Klasse x Standardachievement Mathematik Mitschüler*innen.

Betrachtet man die in Abbildung 1 dargestellten Mittelwerte, zeigen sich nur sehr geringfügige Unterschiede der Mittelwerte zwischen Klassen ohne Schüler*innen mit diagnostiziertem SPF und Klassen, in denen jeweils zwischen einer und fünf Schüler*innen mit SPF anwesend sind. So beträgt der Mittelwertsunterschied bei der Standarderreichung für Mathematik in Klassen ohne Schüler*innen mit SPF 537 Punkte, bei Klassen mit einer Schüler*in mit SPF 531 Punkte und bei Klassen mit 4 Schüler*innen ebenfalls 531 Punkte. Für Deutsch ergibt sich ein relativ identes Bild der Verteilung der Klassenmittel-

werte nach Anzahl der Schüler*innen mit SPF in der Klasse, auf die Darstellung wird daher hier verzichtet.

In Tabelle 1 sind die Koeffizienten der Regressionsmodelle für den Mathematikdatensatz dargestellt. Auf Angabe der Signifikanzniveaus wurde verzichtet, da es sich um einen Volldatensatz handelt.

Tab. 1: Koeffizienten Mathematikdatensatz

	Leermodell	Modell 2	Modell 3	Modell 4	Modell 5	Modell 6	Modell 7
Individualebene							
Geschlecht		18,22		18,05	3,59	3,6	3,6
Alter		-29,35		-28,51	-13,27	-13,28	-13,30
Ausbildung Eltern		5,18		5,25	0,63	0,63	0,63
Bücher Haushalt		15,49		14,91	2,69	2,69	2,69
HISEI		0,84		0,82	0,24	0,25	0,25
Klassenebene							
Anzahl Schüler*innen SPF			-1,29	-0,73	-0,20	1,18	3,82
Erstsprache n, Deutsch			-3,72	-2,44	-1,09	-1,09	-1,09
R2-Level 2	0	0,23	0,26	0,44	0,32	0,32	0,32
R2-Level 1	0	0,17	0	0,17	0,03	0,03	0,03
ICC's	0,18	0,14	0,14	0,12	0,07	0,07	0,07

Insgesamt erklären die Modelle hinsichtlich des Mathematikachievements bis zu 17% der Varianz in der Mathematikleistung auf der Individualebene und bis zu 44% der Varianz auf Klassenebene. Mit einem Intra-class Korrelationskoeffizient von 0,18 können 18% der Varianz durch Klassenebeneneffekte (between class) erklärt werden und 83% durch Prädiktoren, die innerhalb einer Klasse wirken (within class effects). Die Varianzerklärung steigt bis Modell vier an (vollständiges Modell am kompletten Datensatz) und

sinkt anschließend bei den Modellen fünf bis sieben wieder ab. Dies ist dadurch erklärbar, dass bei den letzten drei Modellen nur Daten von Schüler*innen aus dem unteren Leistungsquartil verwendet wurden und dadurch der verwendete Datensatz in Umfang und Varianz deutlich eingeschränkt ist.

In Tabelle 2 sind die Koeffizienten für den Deutschdatensatz abgebildet. Auf die Angaben des Signifikanzniveaus wurde wiederum verzichtet.

Tab. 2: Koeffizienten Deutschdatensatz

	Leermodell	Modell 8	Modell 9	Modell 10	Modell 11
Individualebene					
Geschlecht (männlich)		-18,65	-19,49	-10,65	-8,1
Alter		-19,59	-18,41	-12,11	1,81
Erstsprache (1: Deutsch; 2: Andere)		26,41	24,99	10,30	7,30
Kindergartenbesuch		11,61	11,46	5,31	4,08
Sozialstatus		14,51	11,43	3,24	6,36
Selbstkonzept		37,01	37,78	11,52	17,08
Klassenebene					
Anzahl SuS SPF			-0,98	0,13	-0,32
Klassengröße			-0,73	-0,34	-0,31
Sozialstatus Klasse			24,53	9,03	9,21
Selbstkonzept Klasse			16,37	3,56	6,37
R ² -Level 2	0	0,37	0,57	0,74	0,55
R ² -Level 1	0	0,35	0,34	0,10	0,11
ICC	0,17	0,13	0,11	0,07	0,05

Ähnlich wie beim Mathematikmodell können mit einem ICC von 0.17 ca. 17 % der Varianz durch Klassenebeneneffekte (between class) erklärt werden und 82 % durch Prädiktoren, die innerhalb einer Klasse wirken (within class effects). Insgesamt erklären die Modelle bis zu 35 % der Varianz in den Standard-Scores für Deutsch auf der Individualebene und bis zu 74 % der Varianz auf Klassenebene. Die höhere Varianzerklärung bei den Deutschmodellen liegt zu einem Großteil an der Berücksichtigung der Variablen Selbstkonzept und auch Sozialstatus auf Klassenebene, die bei den Mathematikmodellen fehlen.

Diskussion

Insgesamt zeigt schon der deskriptive Vergleich der Klassenmittelwerte für Mathematik nur sehr geringe Unterschiede bei der Leistungserreichung von Schüler*innen in Klassen mit Kindern mit SPF und Klassen ohne Kinder mit SPF (6 Punkte Unterschied zwischen Klassen mit Kindern mit SPF und Klassen ohne Kinder mit SPF für Mathematik). Zudem ist, auf Basis der Daten, auch nicht von einem linearen Zusammenhang auszugehen, also dass ein Mehr an Kindern mit SPF zu einer stetigen Verschlechterung/Verbesserung der Standarderreichung der Mitschüler*innen führt. Diese Ergebnisse werden auch von den durchgeführten Mehrebenenmodellen gestützt.

Auch die Ergebnisse der Mehrebenenregression zeigen nur geringe Effekte der Anzahl der Schüler*innen mit SPF auf das Abschneiden ihrer Kolleg*innen bei der BIST-Überprüfung sowohl für Deutsch als auch für Mathematik. Für die BIST-Überprüfung in Mathematik reichen die Beta-Koeffizienten dabei von $-1,29$ – $3,82$ und für die BIST-Überprüfung in Deutsch von $-0,98$ – $0,13$ für die einzelnen Modelle. Das bedeutet, zum Beispiel für Modell 4 bei der Mathematikstandardüberprüfung, dass die Erhöhung der Anzahl der Schüler*innen mit SPF im Schnitt zu einer Verringerung der Mathematikleistung der übrigen Schüler*innen in einer Klasse um $-0,73$ Punkte führt. Setzt man dies in Beziehung zur Verteilung der Mathematikleistung ($M=533$; $SD: 99,89$), ist davon auszugehen, dass es sich dabei nur um einen sehr geringen Effekt handelt, der in der Praxis kaum von Bedeutung sein wird. Ein sehr ähnliches Bild zeigt sich dabei für Deutsch, wo etwa bei Modell neun die Erhöhung der Anzahl der Schüler*innen mit SPF zu einer Verringerung der Standard-Scores ihrer Mitschüler*innen von ca. einem Punkt führt. Bei einem Mittelwert von 523 Punkten ($SD=100,48$) kann hier ebenfalls von einer vernachlässigbaren Effektstärke ausgegangen werden.

Gleichzeitig kann je nach Modell die Anwesenheit von Schüler*innen mit SPF auch zu einem positiven Effekt auf die Leistungserreichung bei der Standardüberprüfung haben. So etwa bei den Modellen sechs und sieben jeweils für das untere Leistungsquartil in Mathematik (Koeffizienten: $1,18$ – $3,82$) oder etwa bei Modell zehn für das untere Leistungsquartil in Deutsch (Beta: $0,13$). Das bedeutet, dass bei diesen Modellen die Erhöhung der Anzahl der Schüler*innen mit diagnostiziertem SPF in einer Klasse sogar zu einer Erhöhung der Standard-Scores bei ihren Mitschüler*innen führt. Damit ist analog zu den

negativen Beta-Koeffizienten davon auszugehen, dass zwar inklusive Beschulungsmodelle einen positiven Effekt auf die Leistungserreichung gerade beim unteren Leistungsquartil aufweisen (sowohl für Leseverständnis als auch Mathematik), dass jedoch die Stärke dieses Effekts ebenfalls in der Praxis vernachlässigbar ist.

Schließlich kann damit auch die Frage nach einem linearen Zusammenhang zwischen Anzahl der Schüler*innen mit SPF und die Leistungserreichung ihrer Mitschüler*innen endgültig verneint werden. Vielmehr scheint es sich um einen nicht-linearen Effekt zu handeln, der je nach betrachteter Population sowohl positive als auch negative Werte annehmen kann. In beiden Fällen dürfte jedoch die Stärke des Effekts in der Praxis keine Rolle spielen. Damit sind die Ergebnisse hinsichtlich der Auswirkungen inklusiver Beschulungsmodelle auf die Leistungserreichung der Schüler*innen ohne SPF in einer Schulklasse sowohl von der Richtung als auch von der Stärke vergleichbar mit jenen in der Literatur bzw. mit jenen in anderen Bildungssystemen (Szumski et al., 2017).

Im Gegensatz dazu weist der sozioökonomische Hintergrund der Schüler*innen, auf individueller Ebene (Mathematik und Deutsch) und als Klassenkompositionsmerkmal (nur für das Leseverständnis), einen von der Stärke her signifikanten Effekt auf, sowohl hinsichtlich der Standarderreichung in Mathematik als auch in Deutsch. Dies zeigt sich auch bei der Varianzerklärung der einzelnen Variablen. Während die Anzahl der Schüler*innen mit SPF gerade für 1 % Varianzerklärung verantwortlich ist, beträgt sie beim sozioökonomischen Status bis zu 17 % (inklusive Selbstkonzept) bei der Standarderreichung Deutsch. Dieses Ergebnis deckt sich mit den Ergebnissen aus der Literatur, für Deutsch (Niklas & Schneider 2013; Philipp, 2011; Rodriguez-Brown, 2011) wie auch für Mathematik (Biedermann et al., 2016; Sirin, 2005). Im Gegensatz dazu weisen die Klassenlevelvariablen „Klassengröße“ (für Deutsch) und „Anzahl der Schüler*innen mit anderer Erstsprache als Deutsch“ (Mathematik) zwar für (fast) alle Modelle negative Beta-Koeffizienten auf, allerdings nur in einer vergleichbaren Größe wie die Anzahl der Schüler*innen mit SPF. In der Praxis dürften daher die Auswirkungen dieser beiden Prädiktoren einen von der Stärke her sehr überschaubaren Effekt auf das Standardachievement der Schüler*innen einer Schulklasse haben.

Hinsichtlich der Auswirkungen des Selbstkonzepts auf die Leseleistungen entsprechen die Ergebnisse, sowohl auf Individual- als auch auf Klassenebene, den Erwartungen. Auf Individualebene führt, je nach Modell, die Erhöhung des Selbstkonzepts um einen Punkt zu einem Anstieg des Leseachievements bei der Standardüberprüfung von 11 bis 37 Punkten. Auch auf Klassenebene ist der Anstieg mit 4–16 Punkten relativ stark ausgeprägt und deutlich stärker als der Zusammenhang zwischen Anzahl der Schüler*innen mit SPF und der Leistungserreichung bei der Standardüberprüfung der übrigen Schüler*innen in der Klasse.

Die Ergebnisse bezüglich der Individualvariablen Geschlecht und Alter entsprechen den Erwartungen. Mädchen zeigen eine deutlich bessere Leseleistungen als Jungen (je nach

Modell zwischen 8 und 20 Punkten Differenz), während Schüler bessere Mathematikleistungen zeigen als Schülerinnen (je nach Modell zwischen 3 und 18 Punkte). Relativ deutliche Effekte ergaben sich zudem auch für das Alter, wobei diese bei der Standardüberprüfung in Deutsch und Mathematik in Stärke und Richtung relativ ident sind. Mit steigendem Lebensalter schneiden die Schüler*innen im Schnitt daher deutlich schlechter bei der Standardüberprüfung ab, als dies jüngere Schüler*innen tun. Für die Modelle für die Standardüberprüfung Deutsch wurde zudem auch die Dauer des Kindergartenbesuchs berücksichtigt. Auch hier wiederum entsprachen die Ergebnisse den Erwartungen (Anders, 2013). Die positiven Beta-Koeffizienten zeigen deutlich eine durchschnittliche Zunahme der Leseleistung, je länger die betreffenden Schüler*innen den Kindergarten besuchten.

Insgesamt zeigen die Ergebnisse dieses Beitrags daher, dass der Einfluss der Anwesenheit von Schüler*innen mit SPF auf die Leistungserreichung ihrer Mitschüler*innen sowohl bei der Standarderreichung in Deutsch als auch in Mathematik je nach Subpopulation sowohl positiv als auch negativ sein kann und zudem in seiner Stärke sehr überschaubar ist bzw. in der Praxis keine Rolle spielt. Die Befürchtung, dass die Anwesenheit von Schüler*innen mit diagnostiziertem SPF negative Auswirkungen auf die Leistungserreichung ihrer Mitschüler*innen nach sich zieht, lässt sich anhand der empirischen Daten – zumindest für die Primarstufe – in keiner Weise bestätigen. Vielmehr ist davon auszugehen, dass inklusive Klassensettings auch auf die Leistungserreichung von bestimmten Subpopulationen eine positive Wirkung aufweisen, auch wenn die Stärke dieses Effekts ebenfalls sehr gering ausfällt.

Literatur

- Anders, Y. 2013. Stichwort: Auswirkungen frühkindlicher Institutioneller Betreuung und Bildung. *Zeitschrift Für Erziehungswissenschaft*, 16, 237–275. <https://www.doi.org/10.1007/s11618-013-0357-5>
- Biedermann, H., Weber, C., Herzog-Punzenberger, B. & Nagel, A. (2016). Auf die Mitschüler/innen kommt es an? Schulische Segregation – Effekte der Schul- und Klassenzusammensetzung in der Primarstufe und der Sekundarstufe. 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). Leykam.
- BIFIE. 2013. Standard Setting Mathematik. Technische Dokumentation – BIST-Ü Mathematik 4. Schulstufe. <https://www.bifie.at/material/ueberpruefung-der-bildungsstandards/technische-dokumentation/>
- Dessemonet, R. S., Bless, G. & Morin, D. (2012). Effects of Inclusion on the Academic Achievement and Adaptive Behaviour of Children with Intellectual Disabilities. *Journal of Intellectual Disability Research*, 56(6), 579–587. <https://doi.org/10.1111/j.1365-2788.2011.01497.x>
- Cosier, M., Causton-Theoharis, J. & Theoharis, G. 2013. Does Access Matter? Time in General Education and Achievement for Students With Disabilities. *Remedial and Special Education*, 34(6), 323–332. <http://www.doi.org/10.1177/0741932513485448>.
- Groeneveld, I., and M. Knigge. 2015. Moderation primärer sozialer Disparitäten im Leseverständnis in Abhängigkeit vom wahrgenommenen Verhalten der Lehrkraft und der Klassenzusammensetzung. *Zeitschrift für Bildungsforschung*, 5(1), 51–72. <http://www.doi.org/10.1007/s35834-014-0115-7>

- Haerberlin, U., Bless, G., Moser, U. & Klaghofer, R. (1991). *Die Integration von Lernbehinderten: Versuche, Theorien, Forschungen, Enttäuschungen, Hoffnungen*. Haupt.
- Hanushek, E. A., Kain, J. F. & Rivkin, S. G. (2002). Inferring Program Effects for Special Populations: Does Special Education Raise Achievement for Students with Disabilities? *Review of Economics and Statistics*, 84(4), 584–599. <http://doi.org/10.1162/003465302760556431>.
- Hienonen, N., Lintuvuori, M., Jahnukainen, M., Hotulainen, R. & Vainikainen, M.-P. (2018). The Effect of Class Composition on Cross-curricular Competences – Students with Special Educational Needs in Regular Classes in Lower Secondary Education. *Learning and Instruction*, 58, 80–87.
- Kalambouka, A., Farrell, P., Dyson, A. & Kaplan, I. (2007). The Impact of Placing Pupils with Special Educational Needs in Mainstream Schools on the Achievement of their Peers. *Educational Research*, 49(4), 365–382. <https://www.doi.org/10.1080/00131880701717222>
- Krammer, M., Seifert, S. & Gasteiger-Klicpera, B. (2021). The presence of students identified as having special needs as a moderating effect on their classmates' reading comprehension scores in relation to other major class composition effects. *Educational Studies*. <https://doi.org/10.1080/03055698.2021.1875320>
- Krammer, M., Gasteiger-Klicpera, B., Holzinger, A. & Wohlhart, D. (2019). Inclusion and standards achievement: the presence of pupils identified as having special needs as a moderating effect on the national mathematics standards achievements of their classmates. *International Journal of Inclusive Education*. <https://doi.org/10.1080/13603116.2019.1573938>
- Kristoffersen, J. H. G., Krægpøth, M. V., Nielsen, H. S. & Simonsen, M. (2015). Disruptive School Peers and Student Outcomes. *Economics of Education Review*, 45, 1–13.
- Niklas, F. & Schneider, W. (2013). Home Literacy Environment and the Beginning of Reading and Spelling. *Contemporary Educational Psychology*, 38, 40–50. <https://www.doi.org/10.1016/j.cedpsych.2012.10.001>
- Philipp, M. (2011). *Lesesozialisation in Kindheit und Jugend. Lesemotivation, Leseverhalten und Lesekompetenz in Familie, Schule und Peer-Beziehungen*. Kohlhammer.
- Reilly, D., Neumann, D. L. & Andrews, G. (2018). Gender Differences in Reading and Writing Achievement: Evidence from the National Assessment of Educational Progress (NAEP). *American Psychologist*. Advance online publication. <http://www.doi.org/10.1037/amp0000356>
- Rodriguez-Brown, F. V. (2011). Family Literacy. A Current View of Research on Parents and Children Learning Together. In M. L. Kamil, P. D. Pearson, E. B. Moje & P. P. Afflerbach (Hrsg.), *Handbook of Reading Research*. Volume IV. (S. 726–753). Routledge.
- Ruijs, N. (2017). The Impact of Special Needs Students on Classmate Performance. *Economics of Education Review*, 58, 15–31. <https://doi.org/10.1016/j.econedurev.2017.03.002>
- Shifrer, D., Muller, C. & Callahan, R. (2011). Disproportionality and Learning Disabilities: Parsing Apart Race, Socioeconomic Status, and Language. *Journal of Learning Disabilities*, 44(3), 246–257. <http://www.doi.org/10.1177/0022219410374236>
- Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, 75(3), 417–453. <https://doi.org/10.3102/00346543075003417>
- Stubbe, T., Schwippert, K. & Wendt, H. (2016). Kapitel 10 – Soziale Disparitäten der Schülerleistungen in Mathematik und Naturwissenschaften. In H. Wendt, W. Bos, C. Selter, O. Köller, K. Schwippert & D. Kasper (Hrsg.), *TIMMS 2015: Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 300–317). Waxmann.
- Suchan, B. & Breit, S. (2016). *Pisa 2015: Grundkompetenzen am Ende der Pflichtschulzeit im internationalen Vergleich*. Leykam.
- Szumski, G., Smogorzewska, J. & Karwowski, M. (2017). Academic Achievement of Students without Special Educational Needs in Inclusive Classrooms: A Meta-analysis. *Educational Research Review*, 21, 33–54. <https://doi.org/10.1016/j.edurev.2017.02.004>
- Wagner, W., Helmke, A., Schrader, F. W., Eichler, W., Thomè, G. & Willenberger, H. (2008). Selbstkonzept und Motivation im Fach Deutsch. In DESI-Konsortium [Hrsg.], *Unterricht und Kompetenzerwerb in Deutsch und Englisch*. Ergebnisse der DESI Studie (S. 231–243). Beltz.